

**НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ УКРАЇНИ**  
**«КИЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ**  
**імені ІГОРЯ СІКОРСЬКОГО»**  
**ФІЗИКО-ТЕХНІЧНИЙ ІНСТИТУТ**  
Кафедра інформаційної безпеки

«На правах рукопису»

УДК 519.876.5

«До захисту допущено»

В.о. завідувача кафедри

\_\_\_\_\_ М.В.Грайворонський

“    ”    \_\_\_\_\_ 2018 р.

**Магістерська дисертація**  
**на здобуття ступеня магістра**

зі спеціальності:    113    Прикладна математика

на тему: Оцінка емоцій як прихованих культурних тенденцій за допомогою соціальних медіа

Виконав (-ла): студент (-ка) 2 курсу, групи ФІ-72мп  
(шифр групи)

Шуплецов Ілля Костянтинович  
(прізвище, ім'я, по батькові)

(підпис) \_\_\_\_\_

Науковий керівник д.т.н, професор, засл. діяч Качинський Анатолій  
Броніславович

(посада, науковий ступінь, вчене звання, прізвище та ініціали)

(підпис)

Консультант \_\_\_\_\_

(назва розділу)

(науковий ступінь, вчене звання, прізвище, ініціали)

(підпис)

Рецензент \_\_\_\_\_

(посада, науковий ступінь, вчене звання, науковий ступінь, прізвище та ініціали)

(підпис)

Засвідчую, що у цій магістерській  
дисертації немає запозичень з праць інших  
авторів без відповідних посилань.

Студент \_\_\_\_\_  
(підпис)

Київ – 2018 року

**НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ УКРАЇНИ**  
**«КИЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ**  
**імені ІГОРЯ СІКОРСЬКОГО»**  
**ФІЗИКО-ТЕХНІЧНИЙ ІНСТИТУТ**  
Кафедра інформаційної безпеки

Рівень вищої освіти – другий (магістерський) за освітньо-професійною програмою  
Спеціальність (спеціалізація) – 113 Прикладна математика («Аналітичні методи безпеки інформації»)

ЗАТВЕРДЖУЮ

В.о. завідувача кафедри

\_\_\_\_\_ М.В.Грайворонський  
(підпис)

« \_\_\_\_\_ » \_\_\_\_\_ 2018 р.

**ЗАВДАННЯ**  
**на магістерську дисертацію студенту**

Шуплєцову Іллі Костянтиновичу  
(прізвище, ім'я, по батькові)

1. Тема дисертації

Оцінка емоцій як прихованих культурних тенденцій за допомогою соціальних медіа

науковий керівник дисертації д.т.н, професор, засл. діяч Качинський Анатолій Броніславович,

затверджені наказом по університету від «15 » листопада 2018 р. № 4171-с

2. Термін подання студентом дисертації \_\_\_\_\_

3. Об'єкт дослідження \_\_\_\_\_  
\_\_\_\_\_

4. Вихідні дані \_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_

5. Перелік завдань, які потрібно розробити \_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_

6. Орієнтовний перелік ілюстративного матеріалу \_\_\_\_\_  
\_\_\_\_\_

7. Орієнтовний перелік публікацій \_\_\_\_\_

8. Консультанти розділів дисертації\*

Розділ	Прізвище, ініціали та посада консультанта	Підпис, дата	
		завдання видав	завдання прийняв

9. Дата видачі завдання \_\_\_\_\_

Календарний план

№ з/п	Назва етапів виконання магістерської дисертації	Термін виконання етапів магістерської дисертації	Примітка

Студент

\_\_\_\_\_ (підпис)

\_\_\_\_\_ (ініціали, прізвище)

Науковий керівник дисертації

\_\_\_\_\_ (підпис)

\_\_\_\_\_ (ініціали, прізвище)

\_\_\_\_\_

\* Консультантом не може бути зазначено наукового керівника магістерської дисертації.

## РЕФЕРАТ

Обсяг роботи 85 сторінки, 21 ілюстрації, 7 таблиць, 2 додатки, 17 джерело літератури.

Мета дослідження полягає в отриманні оцінки емоційного стану суспільства використовуючи соціальні мережі.

Об'єкт дослідження є вплив емоцій на геополітичний вибір суспільства.

Предмет дослідження є оцінка впливу ЗМІ на емоційний стан суспільства в соціальних мережах.

ГЕОПОЛІТИКА, ЕМОЦІЇ, ГЕОПОЛІТИКА ЕМОЦІЙ, НЕЙРОННІ МЕРЕЖІ, ЧАСОВІ РЯДИ, ГЛОБАЛІЗАЦІЯ, АНАЛІЗ ДАНИХ, TWITTER, TWITTER API.



## **ABSTRACT**

The volume of work is 85 pages, 21 illustrations, 7 tables, 2 annexes, 17 sources of literature.

The purpose of the study is to obtain an assessment of the emotional state of society using social networks.

The object of research is the influence of emotions on the geopolitical choice of society.

The subject of the study is an assessment of the influence of the media on the emotional state of society in social networks.

GEOPOLITICS, EMOTIONS, EMOTION GEOPOLITICS, NEURAL NETWORKS, TIME RANGES, GLOBALIZATION, DATA ANALYSIS, TWITTER, TWITTER API.

## ЗМІСТ

Перелік умовних позначень, символів, одиниць, скорочень і термінів .....	8
Вступ.....	9
1 Геополітика та емоції.....	11
1.1 Емоції.....	11
1.2 Геополітика.....	14
1.3 Глобалізація .....	22
1.4 Геополітика емоцій .....	25
1.5 Емоція надії .....	27
1.6 Емоція страху .....	28
1.7 Емоція приниження.....	28
Висновки до розділу 1 .....	29
2 Теоретичні відомості .....	30
2.1 Twitter API.....	30
2.2 Аналіз даних .....	34
2.3 Машинне навчання .....	35
2.4 Нейронні мережі.....	41
2.5 Часові ряди .....	56
Висновки до розділу 2 .....	63
3 Практична частина .....	64
3.1 Збір даних.....	64
3.2 Аналіз даних .....	68
3.3 Результати.....	70
3.3.1 Результат групи 112 News.....	70
3.3.2 Результат групи UnianInfo .....	72
3.3.3 Результат групи UaNews .....	74
3.3.4 Результат групи kanalukraine .....	76
3.3.5 Результат групи News On Ukraine .....	78
3.3.6 Результат за 5 групами .....	80

Висновки до розділу 3 .....	81
Висновки .....	82
Перелік джерел посилання .....	84
ДОДАТКИ.....	86
ДОДАТОК А.....	87
ДОДАТОК В.....	92

**ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ, СИМВОЛІВ, ОДИНИЦЬ,  
СКОРОЧЕНЬ І ТЕРМІНІВ**

API – application programming interface

API – прикладний програмний інтерфейс

ЗМІ – засоби масової інформації

## ВСТУП

**Актуальність роботи.** Проблема геополітики відноситься до життєво важливих для України у практично-політичному аспекті. Україна складає вагомий геополітичний чинник у загальносвітовому масштабі.

Глобалізація - процес всесвітньої економічної, політичної та культурної інтеграції та уніфікації. У широкому розумінні - перетворення певного явища на планетарне, таке, що стосується всієї землі. Основними наслідками глобалізації є міжнародний поділ праці, міграція в масштабах усієї планети капіталу, людських та виробничих ресурсів, стандартизація законодавства, економічних та технічних процесів, а також зближення культур різних країн.

В епоху глобалізації емоції стали невід'ємною частиною розуміння всієї складності світу, в якому ми живемо. Посилюючись ЗМІ, вони одночасно і відображають глобалізацію, і реагують на неї, в свою чергу впливаючи на геополітику.

**Мета та завдання дослідження.** Мета дослідження полягає в отриманні оцінки емоційного стану суспільства використовуючи соціальні мережі.

Завдання дослідження полягає в:

- завантажити дані з соціальних медіа
- отриманні данні розбити на рівні інтервали по часу.
- Класифікувати данні за трьома класами, а саме надія, страх та приниження.
- Після цього по класифікованим даним побудувати часові ряди.

**Об'єкт дослідження** є вплив емоцій на геополітичний вибір суспільства.

**Предмет дослідження** є оцінка впливу ЗМІ на емоційний стан суспільства в соціальних мережах.

**Методи дослідження,** які використовуються в роботі, ґрунтуються на використанні прикладних програмних інтерфейсів для збору даних з

соціальних мереж. Для класифікації даних було використано машинне навчання, а саме нейронні мережі. Та були використані часові ряди.

**Наукова новизна** отриманих результатів полягає в отриманій оцінці впливу ЗМІ на емоційний стан суспільства.

# 1 ГЕОПОЛІТИКА ТА ЕМОЦІЇ

## 1.1 Емоції

Емоція (від лат. Emovere - вражаю, хвилюю) - психічний процес середньої тривалості, що відображає суб'єктивне оціночне ставлення до існуючих або можливих ситуацій і об'єктивного світу[3].

Емоції - це суб'єктивні реакції людини на впливи зовнішніх і внутрішніх подразників, що відображають у формі переживань їх особисту значимість для суб'єкта і проявляються у вигляді задоволення або незадоволення.

Емоції характеризуються трьома компонентами: пережитим або усвідомленими в психіці відчуттям емоції; процесами, що відбуваються в нервовій, ендокринній, дихальній, травній та інших системах організму; спостерігаються виразними комплексами емоцій, в тому числі, на обличчі. Емоції відрізняють від інших видів емоційних процесів: афектів, почуттів і настроїв. Емоції, як і багато інших психічні явища, вивчені слабо і розуміються різними авторами по-різному, тому вищенаведене визначення не можна вважати ні точним, ні загальноприйнятим.

У вузькому значенні слова емоції - це безпосереднє, тимчасове переживання якогось почуття. Так, якщо розглянути почуття, які відчують уболівальниками, на трибуні стадіону і спорту взагалі (почуття любові до футболу, хокею, тенісу), то ці переживання не можна називати емоцією. Емоції тут будуть представлені станом насолоди, захоплення, які відчуває уболівальник, спостерігаючи хорошу гру.

За емоціями була визнана важлива позитивна роль в житті людей, і з ними стали пов'язувати такі позитивні функції: мотиваційно-регулююча, комунікативну, сигнальну і захисну.

Мотиваційно-регулююча функція полягає в тому, що емоції беруть участь у мотивації поведінки людини, можуть спонукати, спрямовувати і регулювати. Іноді емоції можуть замінювати собою мислення в регуляції поведінки.

Комунікативна функція полягає в тому, що емоції, точніше, способи їх зовнішнього вираження, несуть в собі інформацію про психічному і фізичному стані людини. Завдяки емоціям ми краще розуміємо один одного. Спостерігаючи за змінами емоційних станів, з'являється можливість судити про те, що відбувається в психіці. Коментар: люди, що належать до різних культур, спроможні безпомилково сприймати і оцінювати багато вираження людської особи, визначати у ній такі емоції, як радість, гнів, печаль, страх, відраза, подив. Це відноситься і до тих народів, які взагалі ніколи не перебували в прямих контактах один з одним.

Сигнальна функція. Життя без емоцій так само неможлива, як і без відчуттів. Емоції, стверджував Ч. Дарвін, виникли в процесі еволюції як засіб, за допомогою якого живі істоти встановлюють значимість тих чи інших умов для задоволення актуальних для них потреб. Емоційно-виразні рухи (міміка, жести, пантоміміка) виконують функцію сигналів про те, в якому стані знаходиться система потреб людини.

Захисна функція виражається в тому, що, виникаючи як миттєва, швидка реакція організму, може захистити людину від небезпек.

Встановлено, що чим складніше організована жива істота, чим більш високий щабель на еволюційних сходах вона займає, тим багатше і різноманітніше гамма емоцій, яку вона здатна переживати.

Характер переживання (задоволення або незадоволення) визначає знак емоцій - позитивні і негативні. З точки зору впливу на діяльність людини емоції діляться на стенічні і астенічні. Стенічні емоції стимулюють діяльність, збільшують енергію і напруга сил людини, спонукають його до вчинків, висловлювань. Крилатий вислів: "готовий гори звернути". І, навпаки, іноді переживання характеризуються своєрідною скутістю, пасивністю, тоді говорять про астенічних емоціях. Тому в залежності від ситуації і індивідуальних особливостей емоції можуть по-різному впливати на поведінку. Так, горе може викликати апатію, бездіяльність у слабкої людини,



в той час як сильна людина подвоює свою енергію, знаходячи розраду в роботі і творчості.

Модальність - основна якісна характеристика емоцій, що визначає їх вид за специфікою і особливої забарвленості переживань. За модальності виділяються три базові емоції: страх, гнів і радість. При всьому різноманітті практично будь-яка емоція є своєрідним вираженням однієї з цих емоцій. Тривожність, занепокоєння, страх, жах представляють собою різні прояви страху; злість, дратівливість, лють - гніву; веселощі, радість, торжество - радості[1].

Виділяють три класи емоційних станів:[4].

1. Настрій – найпоширеніший емоційний стан, що характеризується слабкою інтенсивністю.
2. Пристрасть – сильне, глибоке, визначене, виразне, тривале, усвідомлене емоційне переживання. Воно активізує діяльність людини, підпорядковує собі всі її думки та вчинки, мобілізує на подолання перешкод, на досягнення поставленої мети.
3. Афект – емоційне переживання, що характеризується різко вираженою інтенсивністю та відносною короткочасністю.

Емоціям притаманні наступні властивості[4].

1. Універсальність
2. Динамічність
3. Домінантність
4. Накопичення
5. Зміцнення
6. Адаптація
7. Упередженість
8. Заразливність
9. Пластичність
10. Утримання в пам'яті
11. Іррадіація

## 12.Перенос

### 1.2 Геополітика

Геополітика - напрям політичної думки, концепція, про контроль над територією, про закономірності розподілу і перерозподілу сфер впливу (центрів сили) різних держав і міждержавних об'єднань. Відноситься до роду суспільно-географічних наук, є частиною політичної географії.

Геополітика являє собою дисципліну, розташовану на перетині двох наук - політології і соціології. Двоїстий характер геополітики, її методу, термінології та інструментарію послужив причиною того, що вона довгий час не могла знайти собі місця серед класичних академічних дисциплін, що майже на сторіччя сповільнило її повноцінну інституціоналізацію.

Розрізняють традиційну геополітику, нову геополітику геоекономіку і новітню геополітику геофілософію. Традиційна геополітика робить акцент на військово-політичну міць держави і домінуючу роль географічних чинників у захопленні чужих територій. Геоекономіка, на відміну від традиційної геополітики, робить акцент на економічній могутності держави. Новітня геополітика, в якій домінує сила духу над військовою і економічною потужністю, сприяє подоланню традиційного географічного та економічного детермінізму за рахунок розширення базисних факторів, що визначають поведінку держав у міжнародних відносинах[16].

Концепція геополітики виникла в кінці XIX - початку XX століття, в перших роботах вживалося вираз «політична географія». Термін «геополітика» ввів в обіг шведський політолог і державознавець Рудольф Челлен під впливом німецького географа Фрідріха Ратцеля, який в 1897 році опублікував книгу «Політична географія». Вперше він вжив термін в 1899 році, але широку популярність він здобув після виходу книги «Держава як організм» 1916. Поряд з Челленом класиками геополітичної науки вважаються британський географ і політик Хелфорд Маккіндер, американський історик

морської стратегії А. Мехен, німецький географ, зачинатель політичної географії Ф. Ратцель, німецький дослідник К. Хаусхофер, американський дослідник міжнародних відносин М. Дж. Спикмен[16].

З другої половини XX століття предметом розгляду геополітики стали перш за все такі явища і поняття стратегічної географії, як холодна війна, військово-стратегічний паритет, пізніше також - глобалізація, багатополлярний світ, а також в широкий ужиток увійшли поняття наддержава, велика, регіональна, ядерна, космічна, економічна, спортивна держава стосовно державам, що виділяється по комплексу або окремої характеристики і мають вплив на інші країни.

Основним об'єктом вивчення геополітики є геополітична структура світу, представлена безліччю територіальних моделей. Дослідження механізмів та форм контролю над територією — одна з основних задач геополітики. Історичним ядром геополітики виступає географія, метою якої насамперед є дослідження прямих і зворотних зв'язків між властивостями території і балансом (суперництвом або співпрацею) світових силових полів. Методологічним ядром геополітики при цьому є «моделювання» на загальнопланетарному рівні, хоча в складі цієї наукової дисципліни існують і регіональні, і локальні розділи, наприклад, дослідження меж, проблем спірних територій, міждержавних конфліктів і т. д. Проте, регіональні і локальні проблеми можуть успішно досліджуватися тільки в контексті зазначеного методологічного ядра, тобто від загального до конкретного.

У наукових колах геополітика передбачає географічний, історичний та соціологічний аналіз питань, пов'язаних з політикою і територіальними структурами на різних рівнях від державного до міжнародного. При цьому розглядаються політичне, економічне і стратегічне значення географії, в залежності від місця розташування, розміру, функції та взаємовідносин місцевостей і ресурсів[16].

Основними геополітичними школами є:

1. Російська, радянська

2. Німецька
3. Японська
4. Англійська
5. Американська

Німецька геополітична школа.

Яскраві представники цієї школи - Ф. Ратцель, Р. Челлен, К. Хаусхофер, К. Шмітт.

Великий внесок у сучасну геополітику зробив найбільший німецький географ Фрідріх Ратцель (1844-1904). Науку, яка покликана досліджувати зв'язок держави і простору він назвав «політичною географією». Основні положення нової науки Ратцель виклав у своїй книзі «Політична географія» (1897). Дві ідеї були покладені їм в основу своїх міркувань.

По-перше, ідея визначальної ролі географічних факторів для розвитку суспільства. «Як би людство не тягнулося до вищих емпіреї, писав учений, ноги його торкаються землі. Цим перш за все обумовлюється необхідність розгляду географічних умов його існування».

Він вважав, що дослідження зовнішніх географічних і внутрішніх історичних чинників розвитку суспільства повинні йти рука об руку, бо тільки так може вийти справжня оцінка нашого предмета.

Друга ідея була їм сприйнята від дарвінівської теорії еволюції. Держава, в його уявленні, є живий організм, що з'єднує властивості народу і землі, подібно всім організмам, що бореться за своє існування. Будучи живим організмом, держава рухається і росте як ціле.

Ратцель сформулював сім основних законів просторового зростання держави.

1. Розмір держави збільшується з розвитком культури нації;
2. Зростання держави супроводжується такими аспектами розвитку, як ідеї, торгівля, місіонерство, активність;
3. Держава розширюється, приєднуючи та поглинаючи політичні" - одиниці меншого значення;

4. Кордон — це периферійний орган держави, через який проявляється її зростання, сила або слабкість та всі зміни в організмі держави. Сильною є та держава, яка здатна підтримувати тісні у зв'язки між власним прикордонням та серцевиною, не допускаючи сепаратизму;

5. Експансія держави полягає у першочерговому загарбанні найважливіших для її розвитку регіонів;

6. Первісний імпульс експансії надходить до примітивних держав від більш розвинутих цивілізацій;

7. Загальна тенденція до асиміляції слабших націй спонукає до подальшого збільшення території в русі, який підживлює сам себе.

Найбільш повного розвитку німецька традиція геополітичного знання знайшла в особі професора географії, генерал-майора Карла Хаусхофера (1869-1946). Виходець з аристократичної баварської родини, син професора політичної економії і генерал-майора. Син К. Хаусхофера Албрехт теж професор географії, впроваджував геополітичні ідеї в середу нацистської ієрархії. Дружина Хаусхофера березня перевела на німецьку мову книгу англійського геополітика Дж. Фейрграва «Географія і світова міць». Династія професорів, інтелігентна родина з трагічною долею. Син брав участь у змові проти Гітлера і за два тижні до кінця війни був страчений у в'язниці Моабит в Берліні. Сам Хаусхофер з дружиною Мартою в 1946р. покінчили життя самогубством. Санкт-Петербурзький економ-географ С. Б. Лавров пише, що офіційна біографія Хаусхофера сильно відрізняється від реальної. Геополітик Хаусхофер був різнобічним дослідником і цікавився широким спектром проблем. Особливий інтерес він виявляв до Далекого Сходу, перш за все до Японії. Його роботи «Японська імперія в її географічному розвитку», «Геополітика Тихого океану», «Світова політика сьогодні», «Геополітика панідей»[14].

Геополітична доктрина К. Хаусхофера базувалася на планетарно-му дуалізмі: протистоянні морських і континентальних держав. Майбутнє Німеччини як великої держави бачилось у створенні "континентального

блоку" або осі Берлін—Москва—Токіо. Базуючись на ідеї великих просторів, К. Хаусхофер дійшов висновку про необхідність поділу світу вздовж меридіанів на глобальні економічні регіони. Цей новий геополітичний поділ дістав назву панрегіоналізму. Виділення панрегіонів мало ґрунтуватися на об'єднанні держав на основі спільних соціально-політичних та економічних проблем Рисунок 1.1.

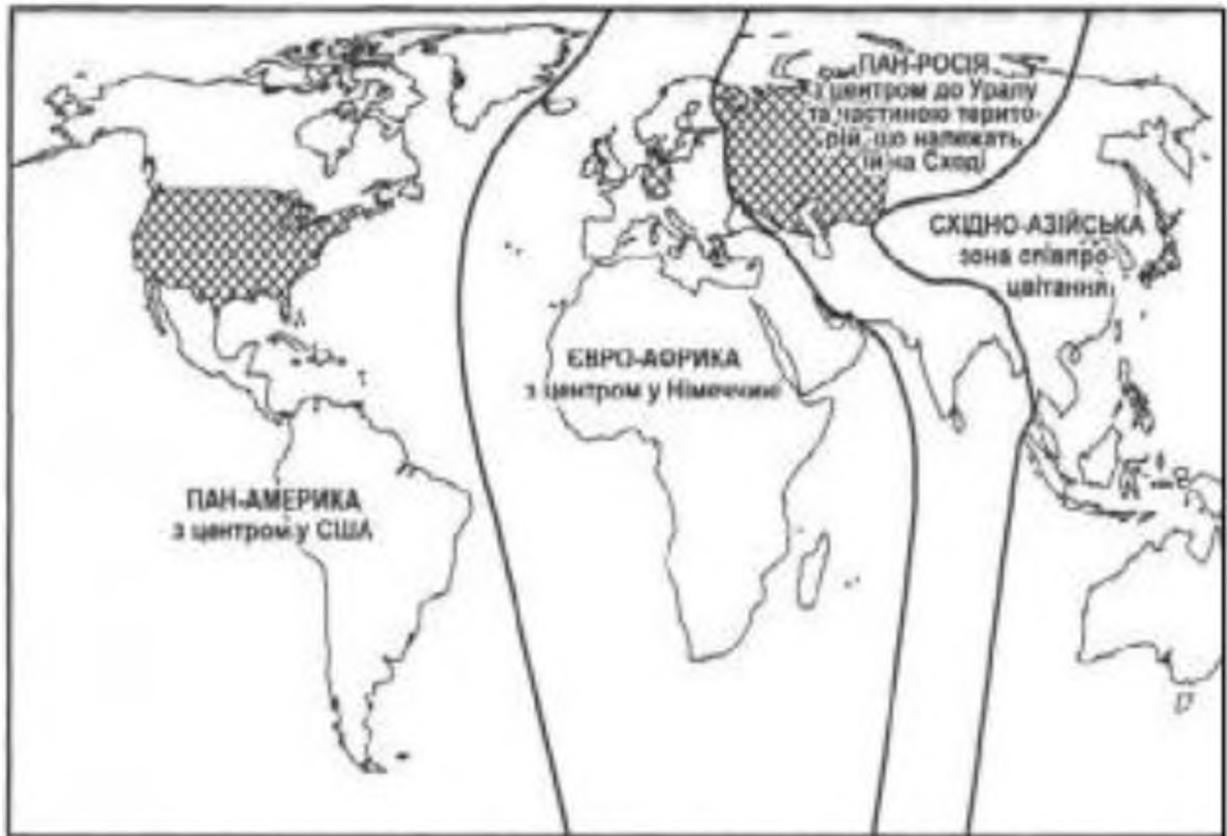


Рисунок 1.1 – Панрегіони К. Хаусхофена

Російська геополітична школа.

У Росії можна відзначити два ідейно-філософські течії, так чи інакше пов'язані з геополітичними інтересами країни - західництво і слов'янофільство. Спочатку 1920-х років в певному шарі російських емігрантів виник ідейно-філософський рух - євразійство. Крім того, в слов'янських країнах, включаючи Росію, в кінці 18-початку 19 століть виник протягом ідейно-політичної думки панславізм. Ще треба зазначити геополітичні ідеї В. П. Семенова Тянь-Шанського[14].

Важливим етапом становлення російської геополітики вважається праця В. П. Семенова Тянь-Шанського(1870—1942) "Про могутнє територіальне

володіння стосовно до Росії. Нарис із політичної географії" (1915), методологічна роль якої полягає у відході від дихотомічної парадигми. Історичний розвиток ішов шляхом синтезу морських і континентальних частин Землі, яскравим прикладом чого є історія Росії. Серед оригінальних ідей: виокремлення трьох форм "територіальних систем політичної могутності (кільцеподібна, клаптикова та "від моря до моря"); з метою збереження цілісності Росії необхідно посунути її культурно-економічний центр до географічного центру, у виділену автором Російську Євразію.

Творцем ідеології російських крайніх правих вважається О. Дугін ("Консервативна революція", 1994, "Основи геополітики", 1997), для якого геополітика — це передовсім метод обґрунтування крайньої правої політичної свідомості та пошуку "третього шляху" для Росії[15].

Американська школа геополітики.

Американська школа є найвагомішою і за кількістю геополітиків, і за чисельністю течій. Найбільший вплив на формування американської військової стратегії та зовнішню політику США протягом усього XX ст. справляли ідеї Альфреда Таєра Мегена[15].

А. Меген, по-перше, виділив 6 критеріїв, які визначають геополітичний статус держави[15]. Серед них:

1. Географічне положення, насамперед відкритість держави морям та доступність морських комунікацій. Протяжність сухопутних кордонів, здатність контролювати стратегічно важливі регіони. Здатність флоту до оборони і війни з ворогом;

2. Морфологія державної території, передусім морських узбережжя, та кількість портів;

3. Протяжність берегової лінії;

4. Кількість населення держави, особливо в контексті можливості будувати та обслуговувати флот;

5. Національний характер, насамперед здібність до торгівлі;

6. Політичний характер правління, від якого залежить спрямування кращих природних та людських ресурсів на створення морської сили.

Послідовник А. Мегена Ніколас Спідмен розглядав геополітику як найважливіший науковий інструмент міжна-родної політики, що дає змогу розробити найефективнішу стратегію безпеки. Важливу роль у формуванні політики він відводив географічним чинникам. Найважливіші праці: "Американська стратегія у світовій політиці: Сполучені Штати і баланс сили" (1942), "Географія світу" (1944). Н. Спідмен запропонував власний варіант геополітичної будови світу Рисунок 1.2.



Рисунок 1.2 – Геополітична модель Н. Спідмена

Британська школа геополітики.

Видатні представники - Х. Маккіндер.

Вагомий внесок у становлення науки геополітики зробив британський географ, родоначальник британської школи Гелфорд Маккіндер (Макіндер) (1861-1947). Йому належить авторство багатьох термінів, коло його інтересів був широкий. Деякий час він викладав географію в Оксфорді і Лондонському університеті, потім очолив Лондонську школу економічних наук, а в 1919-20



роках був призначений верховним британським комісаром по півдню Росії[14].

Згідно з Х. Маккіндером, світова історія є результатом конфронтації континентальних та океанічних держав. Схема Х. Маккіндера базується на ієрархії планетарного простору через систему концентричних кіл Рисунок 1.3.



Рисунок 1.3 – Геополітична модель Х. Маккіндера

Важливим внеском Х. Маккіндера у розвиток науки став запропонований ним поділ геополітичної історії світу на:

1. Доколумбовий
2. Колумбовий
3. Постколумбовий

Загалом уся наукова діяльність ученого була підпорядкована висвітленню взаємозалежності просторових відносин та історичної зумовленості: історія людства є частиною життя світового організму.

### 1.3 Глобалізація

Глобалізація - це історичний процес перетворення світу в єдину систему, що володіє єдиним і характеристикам і. Історично таку систему породжує дію чинників глобалізації:

1. Електронні засоби комунікації, здатні стискати до мінімуму розділяють людей час і простір;
2. Технологічні зміни, що дозволяють поширювати по всьому світу вироблену продукцію;
3. Формування глобальних ідеологій, таких, як екологічне або правозахисний рух.

В кінці 90-х рр. минулого століття в науковий обіг увійшов термін «глобалізація». Під глобалізацією розуміють процес об'єднання економік країн світу. створення єдиного правового, економічного та інформаційного простору.

Характерними рисами глобалізації є:

Посилення ролі транснаціональні компанії. Транснаціональні компанії - господарські підприємства, що діють в двох або більше країнах, і в операціях яких міститься певний закордонний компонент. Як правило, транснаціональні компанії є національними за капіталом і міжнародними за місцем докладання. В даний час в світі налічується понад 7 тис. транснаціональні компанії, в сферу діяльності яких входять близько 700 тис. Дочірніх підприємств. Основними країнами базування майже всіх найбільших транснаціональні компанії є США, Великобританія, Японія.

Створення комп'ютерної мережі Інтернет. Інтернет - всесвітня комп'ютерна мережа, створена в США як ресурс Міноборони в 1969 р і розсекречена в загальносвітовий доступ на початку 1990-х рр. Дає можливість:

1. Миттєво вступати в контакт з будь-яким користувачем в кожній точці земної кулі, проводити інтернет-конференції, наради в режимі реального часу за допомогою мультимедійного режимі;

2. Отримувати доступ до необмеженого обсягу інформації;
3. спілкуватися з будь-якими референтними групами і представниками різних еліт, тим самим ліквідовуючи фактор статусних відмінностей;
4. Використовувати як засіб дозвілля і розваги;
5. Виступати в якості навчального інструменту;
6. Здійснювати ведення бізнесу;
7. Рекламувати товари і послуги;
8. Забезпечувати мобільний зв'язок і комфортність існування;
9. Значно підвищувати ефективність взаємодії громадян і уряду;
10. Організовувати ЗМІ в режимі реального часу.

Зростання мегаполісів. Мегаполіси - великі міста з розвиненою інфраструктурою. налічують кілька мільйонів жителів. Вони можуть вважатися новими формами урбанізації.

Мегаполіси для країн є:

1. Посередниками вступу в глобальний простір;
2. Центрами економічного, технологічного, політичного, культурного розвитку.

Стандартизація та уніфікація. Важливими ознаками глобалізації є: стандартизація промислової продукції, уніфікація ведення звітності, документації, міжнародного законодавства.

Стандартизація - звід правил, призначених для забезпечення співробітництва та однозначного тлумачення умов виробництва, комерційної діяльності, засобів контролю і вимог безпеки.

Поряд з розробкою системи світових стандартів здійснюється обов'язкова сертифікація товарів і послуг.

Сертифікат - це документ, що підтверджує якість товару або послуги.

Стандартизація рухається в двох напрямках: кодування товарів і уніфікація товарів і послуг. При цьому штрих-код є глобальним мовою торгівлі.

Маніпуляція масовою свідомістю.

Маніпуляція масовою свідомістю - різновид психологічного впливу, результатом якого є приховане порушення і поява у адресата намірів, не збігаються з його дійсними бажаннями.

Маніпуляція створюється шляхом: використання стереотипів поведінки; підміни понять; створення метафор; міфологізації; відволікання уваги.

З розвитком засобів масової комунікації з'явилася можливість впливати на свідомість мас усього світу. Технології зміни масової свідомості називаються high-hume. Вони спираються на такі засоби, як: телебачення, кіно, масове мистецтво, глобальна мережа.

Телебачення здатне формувати свідомість або створювати у глядача необхідну реальність, яка не має зв'язку з дійсністю.

XX ст. породив специфічний феномен, названий іспанським філософом Х. Ортегою-і-Гассегом «Повстання мас», суть якого полягає в тому, що на зміну щодо упорядкованого життя суспільства приходить його масовізація. В такому масовому суспільстві більшість населення приймає стандарти, пристрасті і стереотипи, нав'язані засобами масової інформації, в результаті чого на історичну арену виходить «людина-маса», який відчуває себе «таким же як всі» і не прагне до самодосконалості. Світ представляється йому широким полем діяльності для застосування своєї енергії і підприємливості.

Саме на хвилях маніпуляції масовою свідомістю, втрати людьми цінностей традиційного характеру, відриві від історичного коріння найбільш довірливі, безпринципні люди можуть встати на шлях екстремізму і тероризму.

Реально результати процесу глобалізації можна бачити в тому, що кордони для економічної, культурної і навіть політичної діяльності стають все більш прозорими, умовними. Транслокальні і транснаціональні мережі - промислові монополії, Інтернет, обмін студентів і професійних кадрів, сфера послуг, організації «зелених» і рух на захист прав людини - сприяють створенню воістину глобальної культурної та економічної системи. В силу дії всіх цих факторів похитнулося традиційне для XX ст. «Геополітичне уяву», яке зводилося до сприйняття світу розділеним на просторові, територіальні та

ідеологічні блоки. Глобалізація тягне за собою формування нової геополітики і нового бачення світу, все більше здобуває образ єдиного цілого, що складається з «потоків» і «мереж».[17]

#### **1.4 Геополітика емоцій**

Хіба геополітика не базується на раціональності, на об'єктивних даних про межі, економічних ресурсах, військової могутності, холодної оцінці політичних інтересів? Тоді як емоції по суті суб'єктивні і абсолютно ірраціональні. Змішувати емоції і геополітику - справа марна, може бути, навіть небезпечне, оскільки воно здатне привести в прірву нерозумності, символом якої служать язичницькі ходи в Нюрнберзі під час поступового занурення Німеччини в варварство під владою Гітлера.

По-перше, не можна повністю зрозуміти світ, в якому ми живемо, не спробувавши виділити і зрозуміти емоції що наповнюють його. І, по-друге, емоції, як холестерин, можуть бути і корисними, і шкідливими. У листопаді 2008 р, по крайній мірі на час, надія переважала над страхом. Стіна расових забобонів впала, як стіна гніту звалилася в Берліні за двадцять років до цього. Очевидно, що перемогу Обама пояснюють об'єктивні, раціональні причини. У звичайному політичному сенсі вона означала неприйняття народом політики попередньої адміністрації, довгої війни і глибокої економічної кризи. І все ж не можна недооцінювати емоційну складову цих виборів і почуття гордості, яке вони породили в багатьох американців.

Точно так само важко зрозуміти російську військову авантюру на Кавказі влітку 2008 р без розгляду її емоційного значення. Сенс послання московського режиму Путіна і президента Медведєва був абсолютно ясний для грузин й інших народів світу: «Імперська Росія повернулася! Після 1989 року ви сміли ставитися до нас поблажливо. Це залишилося в минулому. Ми

готові здолати наше пострадянське приниження і побудувати нову надію на фундаменті вашого страху ».

Тим же влітку 2008 р інший режим прагнув подолати приниження минулого, але не за рахунок військової авантюри, а завдяки міжнародним спортивним досягненням. Беручи у себе Олімпійські ігри, Китай символічно і емоційно повернув країні історичне центральне становище в світі і легітимність на міжнародній арені. За рахунок пишноти церемонії відкриття, архітектурної краси головного стадіону і великої кількості медалей, виграних китайськими спортсменами, Китай здав вступний іспит, який допускає в сучасне життя, і досяг нової вершини надії, що живиться стрімким економічним зростанням.

Але в той час як Китай хапається за надію, арабський світ тоне в стражданнях і негативних емоціях приниження. Не всі араби - і навіть не більшість з них - підтримують ірраціональну і ненависницького доктрину жорстокої «священної війни» із Заходом. Однак навіть багато поміркованих араби відкидають поняття мирних змін і активної громадянської позиції, вважаючи, що всі політичні лідери безчесні і корумповані. Таку реакцію можна зрозуміти, однак вона відображає і посилює почуття відчаю, що обмежує можливості прогресу в усьому арабському світі.

В епоху глобалізації, коли все і всі взаємопов'язані, надзвичайно важливо встановити власну індивідуальність. У світі, в якому панує ідентичність, нас уже визначають не стільки наші політичні переконання і ідеали, скільки наше сприйняття власної суті, впевненість, яку дають нам наші досягнення, і повагу, яке інші відчують або не відчують до нас. У сприйнятті нашої суті величезну роль грають емоції, вони пов'язані з тим, як ми дивимося на інших, - так само як і з тим, як інші дивляться на нас.

Емоції - це одночасно і зображення в дзеркалі, і очей людини, який бачить цю картинку. Емоції викликають зворотну реакцію, що демонструється, наприклад, поведінкою сучасних освічених мусульманок, що вважають за краще на Заході носити хустку, тим самим викликаючи потік зустрічних

емоцій щодо їх ідентичності і мотивів. Ви боїтеся когось, хтось вас принижує, або в вас будить надію чийсь успіх. Таке переплетення взаємозалежних емоцій має вирішальне значення для розуміння світу, в якому панує ідентичність. Страх, приниження і надію можна, таким чином, розглядати як цілком природні і життєво важливі компоненти існування людських істот, як і три складові частини крові: еритроцити, лейкоцити і тромбоцити. Нам необхідні всі три елементи, щоб бути здоровими. Однак здоров'я залежить від рівноваги між ними. Підвищення або зниження кількості будь-який з цих трьох складових частин небезпечно для загальної рівноваги організму і його здоров'я в довгостроковій перспективі. Рівновага емоцій має настільки ж важливе значення для «здоров'я світу», як і «збалансована» кров для здоров'я людей[2].

Страх протистоїть надії, надія - приниження, приниження веде до прямої ірраціональності і іноді навіть насильства: неможливо зрозуміти світ, в якому ми живемо, не розібравшись в емоціях, які беруть участь в його формуванні[2].

### **1.5 Емоція надії**

Надія або сподівання — емоційне переживання, відчуття, емоція, що виникає при очікуванні суб'єктом бажаної події, передчуття такої події. Інше означення сподівання — віра в позитивне, з точки зору суб'єкта, розв'язання ситуації, що склалась у його житті[5].

В українській мові слова надія та сподівання здебільшого використовуються як синоніми. Надія інколи означає цілковите покладання на свою віру, тоді як сподівання має відтінок більшої впевненості. Однак, таке розділення не є певним. Великі сподівання можуть бути великими ілюзіями. Антонімом до слова надія є безнадія[5].

Надія - це протилежність смирення, форма впевненості, яка змушує нас рухатися у напрямку до інших, не побоюючись, що вони сильно від нас відрізняються.

Сьогодні надія пов'язана з наділенням людини економічної і суспільної владою, і вона знаходиться в основному на Сході. Більш того, для значної кількості азіатів важливі не тільки прагнення наздогнати Захід, а й впевненість у своїй здатності зробити це. Якщо віра - це «надія на духовне», азіатський світ стійко рухається за межі віри, сподіваючись на матеріальний прогрес, який можна побачити, почути, помацати, спробувати і випробувати в стрімко мінливому навколишньому світі[2].

### **1.6 Емоція страху**

Страх — внутрішній стан, що обумовлений загрозою реального або передбачуваного лиха. З точки зору психології вважається негативно забарвленим емоційним процесом[6].

Страх - це не тільки емоційний індикатор, це багатогранна реальність. Зрозуміло, фактор страху є невід'ємною частиною захисту від надмірної самовпевненості. Страх - це сила, що дозволяє виживати в світі, повному природних небезпек. Заєць, який не боїться мисливця, не проживе довго. Страх змушує уважніше ставитися до навколишнього світу - це конструктивне застереження, природний захисний інстинкт.

Страх може також стояти біля витоків надії. Після Другої світової війни страх перед новою війною між Францією і Німеччиною став вирішальним фактором в освіті Європейського союзу. А страх перед наслідками глобального потепління може змусити людство вжити заходів, необхідних для захисту нашої планети від біологічної катастрофи.

### **1.7 Емоція приниження**

Приниження - це безсилля, відчуття, породжене насамперед відчуттям того, що ви більше не здатні спільно, як нація чи як релігійна громада, або



особисто, як індивідуум, управляти звий життям. Приниження досягає найвищої точки, коли ви переконані в тому, що Інший вторгся в приватну сферу вашого життя, зробивши вас повністю залежним. У припиненні укладено почуття відчуження по відношенню до цього і ще більше по відношенню до майбутнього - майбутнього, цілком протилежного ідеалізованому славного минулого, майбутнього, в якому ваші політичні, економічні, соціальні та культурні умови продиктовані Іншим.

В якійсь мірі припинення існує у всіх культурах і суспільствах. Як і холестерин, воно і корисно, і шкідливо. Якась ступінь припинення може стати спонукою до соціальному зростанню за рахунок важкої праці: «Я доведу, на що я здатний». «Я вам покажу, що ви були неправі, коли не приймали мене всерйоз». «Я досягну успіху на славу моїх не користувалися повагою і позбавлених прав предків».

### **Висновки до розділу 1**

Емоції мають значення. Вони впливають на настрої народів, на відносини між культурами і на поведінку націй. Ні політичні лідери, ні історики, ні звичайні зацікавлені люди не можуть собі дозволити їх ігнорувати.

## 2 ТЕОРЕТИЧНІ ВІДОМОСТІ

### 2.1 Twitter API

Соціальна мережа Twitter одна з найвідоміших соціальних мереж, яка є дуже популярна в світі. В Twitter інформація йде короткими порціями та вона обмежена 140 символами.

Традиційні засоби масової інформації виходять в соціальні мережі, задля розширення власної аудиторії.

Соціальна мережа Twitter пропонує серію прикладних програмних інтерфейсів для доступу к даним Twitter. Включаючи читання твітів, завантаження інформації з профілю користувачів та відправку повідомлень від ім'я користувача.

Після реєстрації додатки в Twitter в вкладці Keys and Access знаходиться інформація необхідна для аутентифікації. Ключ користувача Consumer Key та секретний ключ користувача Consumer Secret. Маркер доступу Access Token та секретний маркер доступу Access Token Secret.

Twitter API обмежує доступ. Такі обмеження встановлюються для кожного користувача, або маркер доступу.

По досягненню встановлених меж замість запитаних даних Twitter поверне повідомлення про помилку. Більш того, якщо продовжити далі посилати запити до API, час очікування, що вимагається для відновлення нормального доступу, збільшиться.

Twitter підтримує декілька API. Для простоти розділимо їх на два класи: звичайний пошуковий REST API і потоковий Streaming API.

REST API дозволяє звертатися тільки до минулого. Взаємодіючи з Twitter через REST API, можна шукати вже опубліковані твіти та доступні для пошуку. Нерідко ці API обмежують кількість твітів які можна отримати. Зазвичай для доступу відкриті дані не більше ніж тижневої давності. Другий аспект REST API полягає в тому, що вони не гарантують повернення всіх

повідомлень, які були опубліковані в Twitter, так як деякі можуть бути недоступні для пошуку або просто ще не були проіндексовані.

Потоковий API заглядає в майбутнє. Після установки з'єднання його можна залишити відкритим і рухатись вперед по лінії часу. Зберігаючи HTTP-з'єднання відкритим можна отримувати всі твіти, відповідні заданим критеріям фільтрації, у міру їх публікації[7].

Твіт – складний об'єкт. В Таблиці 2.1 перерахунок всі його атрибути та дає короткий їх опис.

Таблиця 2.1 – Атрибути твіту та їх опис

Назва атрибута	Опис
author	Автор твіту
contributors	список учасників
coordinates	Словник з координатами
created_at	час створення твіту
entities	Словник з URL-адресами, хештегами і згадками в твітах
favorite_count	Число користувачів, яким сподобався твіт
favorited	Зазначає, чи сподобався твіт аутентифікованому користувачеві
geo	координати
id	Унікальний ідентифікатор твіту у вигляді числа
id_str	Унікальний ідентифікатор твіту у вигляді рядка
in_reply_to_screen_name	Екранне ім'я твіти
in_reply_to_status_id	Ідентифікатор твіту у вигляді числа
in_reply_to_status_id_str	Ідентифікатор твіту у вигляді строкового значення
in_reply_to_user_id	Ідентифікатор користувача, який залишив твіт, у вигляді числа
in_reply_to_user_id_str	Ідентифікатор користувача, який залишив твіт, у вигляді строкового значення
in_quote_status	Ознака: чи є твіт цитатою
lang	Рядків з кодом мови твіти

Продовження таблиця 2.1

Назва атрибута	Опис
place	Місце, прикріплене до твіту
possibly_sensitive	Ознака наявності URL-адреси в твіті
retweet_count	Скільки разів це повідомлення було ретвітнуте
retweeted	Чи є це повідомлення ретвітом
source	Рядок, що описує інструмент, який використовується для відправки статусу
text	Рядок з текстом повідомлення
truncated	Ознака усічення твіту
user	автор твіту

Далі продемонстровано приклад твіту в форматі JSON.

```
{
  "created_at":
  "entities": {
    "hashtags": {
      "indices": []
      "text":
    }
  }
  "media": {
    "display_url":
    "id":
    "id_str":
    "indices":
    "type":
    "url":
  }
  "symbols":
  "urls":
  "user_mentions": {
    "id":
```

```

    "id_str":
    "indices":
    "name":
    "screen_name":
  }
}
"favorite_count":
"favorited":
"geo":
"id":
"id_str":
"lang":
"retweet_count":
"retweeted":
"text":

```

Атрибут user – це словник представляє користувача, який відправив твіт. Це складний об'єкт з інформацією про користувача, вставлений в твіт.

```

"user": {
    "created_at":
    "description":
    "entities": {
        "description":
    }
    "url":
    "favorites_count":
    "followers_count":
    "following":
    "friends_count":
    "id":
    "id_str":

```

```

    "lang":
    "location":
    "name":
    "screen_name":
    "statuses_count":
    "time_zone":
    "url":
    "utc_offset":
    "verified":
  }

```

Цей приклад демонструє два цікавих аспекти, які потрібно враховувати при аналізі твітів:

1. Сутності мають мітки
2. Профіль користувача повністю вбудований в твіт.

Перший пункт означає спрощення аналізу сутностей, оскільки позбавляє від необхідності явно шукати такі сутності як хештеги, згадки користувачів, вбудовані зображення, тому що всі вони передаються Twitter API.

Другий пункт означає відсутність необхідності зберігати інформацію з профілю користувача десь в іншому місці. Профіль користувача надлишково повторюється в кожному твіті.

## 2.2 Аналіз даних

Аналіз даних - область математики та інформатики, що займається побудовою і дослідженням найбільш загальних математичних методів і обчислювальних алгоритмів вилучення знань з експериментальних (в широкому сенсі) даних; процес дослідження, фільтрації, перетворення і моделювання даних з метою отримання корисної інформації та прийняття

рішень. Аналіз даних має безліч аспектів і підходів, охоплює різні методи в різних областях науки і діяльності.

Інтелектуальний аналіз даних - це особливий метод аналізу даних, який фокусується на моделюванні і відкритті даних, а не на їх описі. Бізнес-аналітика охоплює аналіз даних, який покладається на агрегацію. У статистичному сенсі деякі поділяють аналіз даних на описову статистику, дослідницький аналіз даних і перевірку статистичних гіпотез. Дослідницький аналіз даних займається відкриттям нових характеристик даних, а перевірка статистичних гіпотез - на підтвердження або спростування існуючих гіпотез. Прогнозний аналіз фокусується на застосуванні статистичних або структурних моделей для передбачення або класифікації, а аналіз тексту застосовує статистичні, лінгвістичні і структурні методи для вилучення і класифікації інформації з текстових джерел належать до неструктурованих даними. Все це різновиди аналізу даних[8].

Розділи аналізу даних:

1. Попередня обробка даних
2. Кореляційний аналіз
3. Дисперсійний аналіз
4. Регресійний аналіз
5. Коваріаційний аналіз
6. Дискримінаційний аналіз
7. Кластерний аналіз
8. Аналіз часових рядів

### **2.3 Машинне навчання**

Машинне навчання - клас методів штучного інтелекту, характерною рисою яких є не пряме рішення задачі, а навчання в процесі застосування рішень безлічі подібних завдань. Для побудови таких методів використовуються засоби математичної статистики, чисельних методів,

методів оптимізації, теорії ймовірностей, теорії графів, різні техніки роботи з даними в цифровій формі.

Розрізняють два типи навчання:

Навчання по прецедентах, або індуктивне навчання, засноване на виявленні емпіричних закономірностей в даних.

Дедуктивне навчання передбачає формалізацію знань експертів і їх перенесення в комп'ютер у вигляді бази знань. Дедуктивне навчання прийнято відносити до області експертних систем, тому терміни машинне навчання і навчання по прецедентах можна вважати синонімами.

Багато методів індуктивного навчання розроблялися як альтернатива класичним статистичним підходам. Багато методів тісно пов'язані з витяганням інформації, інтелектуальним аналізом даних.

Як правило, завдання буде полягати в тому, щоб використовувати існуючі дані для розробки моделей, які можна застосовувати для передбачення різних результатів щодо нових даних, як наприклад, передбачення:

1. Чи є повідомлення спамом чи ні.
2. Чи є транзакція по кредитній карті шахрайської.
3. Чи є транзакція по кредитній карті шахрайської.
4. Яка футбольна команда виграє.

Існує кілька типів моделей. Модель навчання з учителем тобто така, де є вибірка, маркована правильними відповідями, на яких проходить навчання. Модель без учителя де правильні відповіді не задані. Так само існують моделі з частковим навчанням де тільки деякі дані марковані правильними відповідями. Моделі з динамічним навчанням де модель повинна постійно пристосовуватися до прибуваючих даних.

При цьому, навіть в найпростішій ситуації зв'язки які нас цікавлять можуть бути описані цілісними сукупностями або сімействами параметризованих моделей, і в більшості випадків ми самі вибираємо таке



сімейство моделей, а потім використовуємо дані, щоб обчислити параметри, які в певній мірі є оптимальними.

Наприклад, можна припустити що зріст людини є приблизно лінійною функцією його ваги, і потім використовувати дані, щоб обчислити вид цієї лінійної функції. Або припустити, що дерево прийняття рішень є хорошим механізмом для діагностики захворювань пацієнтів, і потім використовувати дані, щоб обчислити подібне дерево.

Загальновідомою пасткою в машинному навчанні є перенавчання, що призводить до прогнозної моделі, яка добре працює на даних з навчальної вибірки, але при цьому погано узагальнює на будь-яких нових даних.

Це явище може бути викликано тим, що модель навчена шуму в даних або навчання розрізняти конкретні входять значення, а не ті чинники, які насправді носять для необхідного результату прогнозуючий характер.

Іншу сторону цього явища являє недонавчання, яке призводить до прогнозної моделі, що не працює навіть на навчальній вибірці. Втім, коли це трапляється, зазвичай роблять висновок, що модель недостатньо хороша, і продовжують шукати більш відповідну.

На Рисунку 2.1 виконана поліноміальна апроксимація вибірки з даних.

Горизонтальна пряма відповідає найбільш підходящому значенням полінома 0-го ступеня. Вона вкрай недостатньо підігнана під навчальну вибірку тобто недонавчена. Найбільш підходяще значення кривої полінома 9-го ступеня підходить рівно через кожну точку навчальної вибірки, але воно надмірне підігнано під вибірку тобто перенавчена - якби довелося додати ще кілька точок, то крива, цілком ймовірно, пройшла б повз на великому від них віддаленні.

А ось лінія полінома 1-го ступеня досягає хорошого рівноваги - вона знаходиться досить близько до кожної точки, і пряма, швидше за все, буде знаходитися близько і до нових точок даних[9].

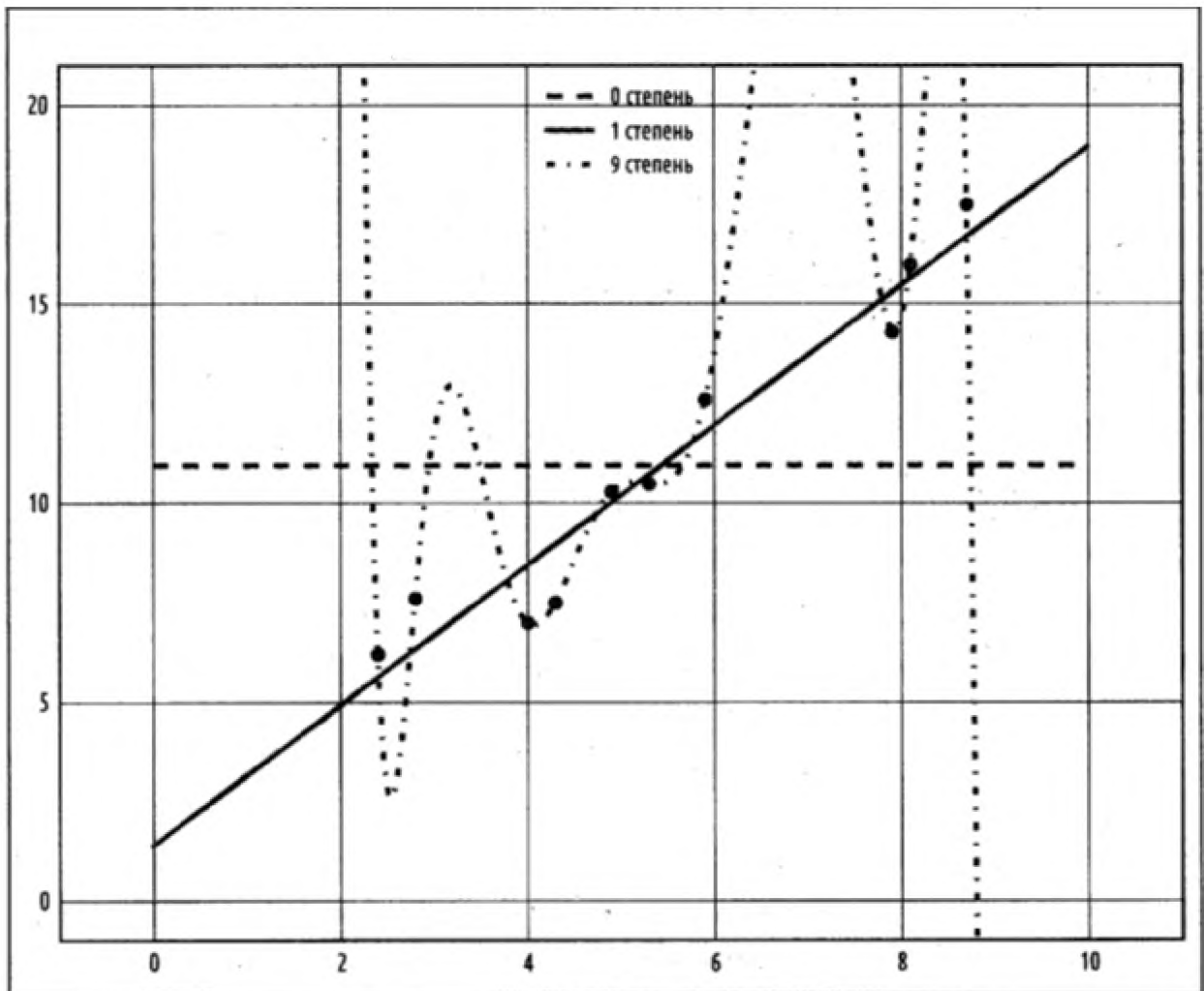


Рисунок 2.1 – Перенавчання та недонавчання

Ясно, що занадто складні моделі призводять до перенавчання і погано узагальнюють за межами даних, на яких вони навчені. Тоді як зробити так, щоб моделі не були занадто складними? Найбільш фундаментальний підхід передбачає використання різних даних для навчання моделі і для перевірки її якості.

Найпростіший спосіб полягає в розбитті набору даних, завдяки чому дві третини використовуються для навчання моделі, після чого продуктивність моделі вимірюється на третини що залишилася.

Якщо модель перенавчена, то можна сподіватися, що вона покаже дуже слабкі результати на контрольних даних. Інакше кажучи, якщо вона добре працює на контрольних даних, то впевненості в тому, що модель навчена ніж перенавчена, набагато вище.

Проте, в кількох випадках це може бути неправильним.

По-перше, в контрольній і навчальній вибірках можуть бути типові схеми, що не узагальнюються на більший набір даних.

По-друге, і це являє значнішу проблему, коли розбивка на контрольну та навчальну вибірки використовується не тільки для того, щоб судити о моделі, але і для того, щоб вибирати серед декількох моделей. У цьому випадку, незважаючи на те, що кожна окрема модель може виявитися неперенавченою, навчальне метаправо, яке змушує вибирати ту модель, яка на контрольній вибірці демонструє найкращий результати, перетворює функцію контрольної вибірки в другу навчальну вибірку.

У такій ситуації необхідно розділити дані на три частини: навчальна підмножина для побудови моделей, валідаційна підмножина для вибору однієї з навчених моделей і контрольна підмножина для оцінювання остаточної моделі.

Ще один спосіб поглянути на проблему перенавчання полягає в пошуку балансу між зміщенням і дисперсією.

Обидва показники дозволяють виміряти результати багаторазового перенавчання моделі на різних навчальних вибірках.

Наприклад, модель з поліномом 0-го ступеня зробить багато помилок практично на будь-якій навчальній вибірці, а значить, має високе зміщення. Однак будь-які дві навчальні вибірки, взяті випадково, дадуть цілком схожі моделі. І тому говорять, що модель має низьку дисперсію. А висока зміщення і низька дисперсія, як правило, відповідає недонавчанню.

З іншого боку, модель з поліномом 9-го ступеня, яка добре підігнана під навчальну вибірку, має дуже низьке зміщення, але дуже високу дисперсію. А це відповідає перенавчанню.

Якщо модель має високе зміщення, то слід додати більше ознак. Перехід моделі з поліномами 0-го ступеня до моделі з поліномами 1-го ступеня в результаті дав велике поліпшення.

Якщо модель має високу дисперсію, то можна точно також видалити ознаки. Втім, ще одне рішення полягає в тому, щоб добути більше даних якщо це можливо.

На Рисунку 2.2 підібраний поліном 9-го ступеня до вибірок різних обсягів. На основі 10 точок даних модель підігнана по всьому полю. Якщо навчити модель на 100 точках даних, то перенавчання набагато знизиться.

Якщо тримати складність моделі постійною, то чим більше даних, тим важче її перенавчити.

З іншого боку, додаткові дані не поліпшить зміщення. Якщо модель не використовує достатньої кількості ознак для пояснення закономірностей в даних, то мобілізація додаткових даних не допоможе.

Ознаки - це будь-які вхідні значення, які передаються в модель.

У найпростішому випадку ознаки просто задані. Якщо потрібно передбачити чийсь зарплату на основі його стажу, то число років і буде єдиною ознакою[9].

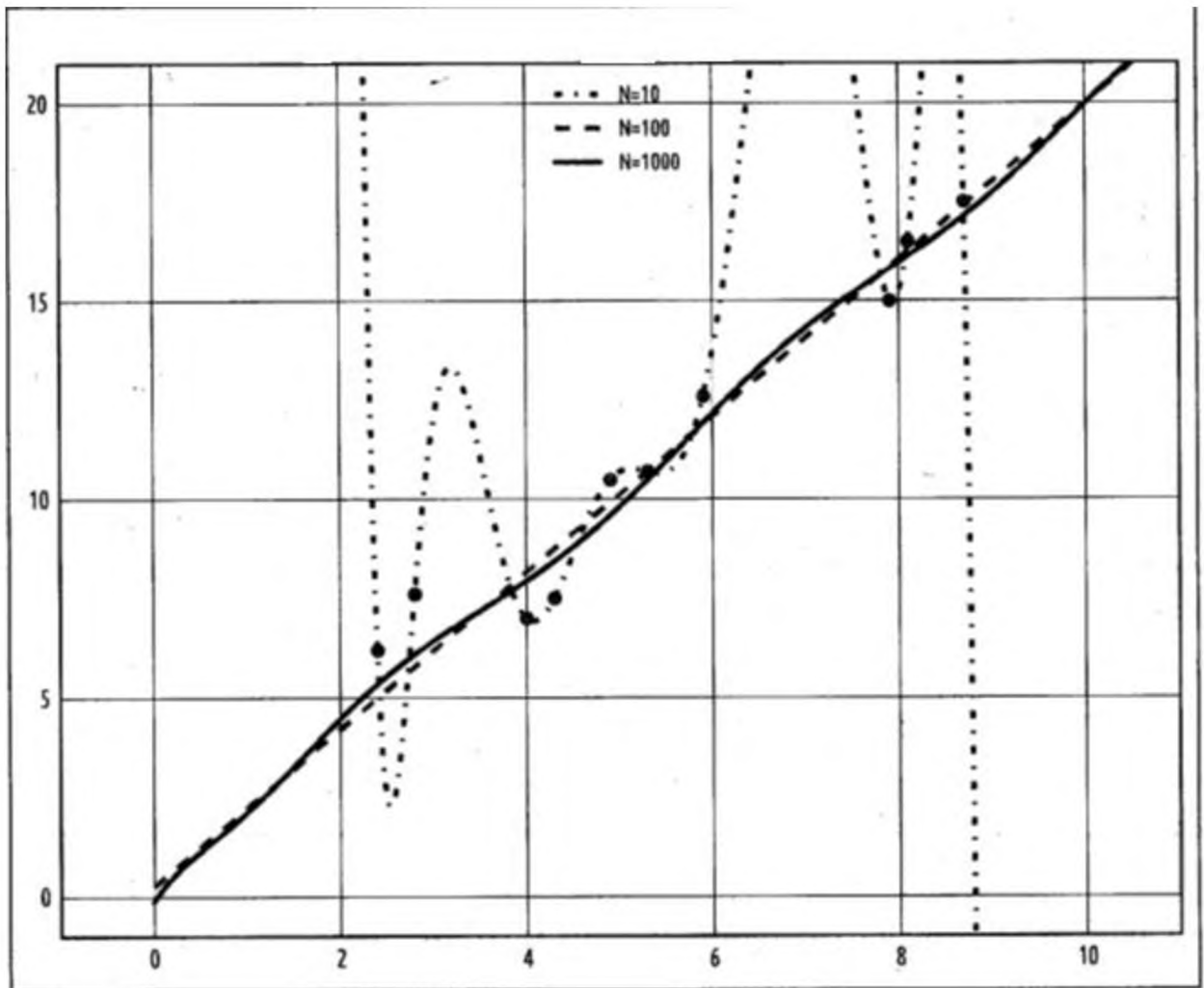


Рисунок 2.2 – зменшення дисперсії

## 2.4 Нейронні мережі

Нейронна мережа - це послідовність нейронів, з'єднаних між собою синапсами. Структура нейронної мережі прийшла в світ програмування прямо з біології. Завдяки такій структурі, машина отримує можливість аналізувати і навіть запам'ятовувати різну інформацію. Нейронні мережі також здатні не тільки аналізувати вхідну інформацію, а й відтворювати її зі своєї пам'яті. Іншими словами, нейронна мережа це машинна інтерпретація мозку людини, в якому знаходяться мільйони нейронів передають інформацію у вигляді електричних імпульсів.

Нейронні мережі використовуються для вирішення складних завдань, які вимагають аналітичних обчислень подібних тим, що робить людський мозок. Найпоширенішими застосуваннями нейронних мереж є:

Класифікація - розподіл даних по параметрах. Наприклад, на вхід дається набір людей і потрібно вирішити, кому з них давати кредит, а кому ні. Цю роботу може зробити нейронна мережа, аналізуючи таку інформацію як: вік, платоспроможність, кредитна історія і тп.

Передбачення - можливість прогнозувати наступний крок. Наприклад, зростання або падіння акцій, ґрунтуючись на ситуації на фондовому ринку.

Розпізнавання - в даний час, саме широке застосування нейронних мереж. Використовується в Google, коли ви шукаєте фото або в камерах телефонів, коли воно визначає положення вашого обличчя і виділяє його і багато іншого.

Тепер, щоб зрозуміти, як же працюють нейронні мережі, давайте поглянемо на її складові і їх параметри.

Нейрон - це обчислювальна одиниця, яка отримує інформацію, виробляє над нею прості обчислення і передає її далі. Вони діляться на три основних типи: вхідний (синій), прихований (червоний) і вихідний (зелений). Також є нейрон зміщення і контекстний нейрон про які ми поговоримо в наступній статті. У тому випадку, коли нейронна мережа складається з великої кількості нейронів, вводять термін шару. Відповідно, є вхідний шар, який отримує інформацію,  $n$  прихованих шарів (зазвичай їх не більше 3), які її обробляють і вихідний шар, який виводить результат. У кожного з нейронів є 2 основні параметри: вхідні дані і вихідні дані. У разі вхідного нейрона:  $input = output$ . В інших, в поле  $input$  потрапляє сумарна інформація всіх нейронів з попереднього шару, після чого, вона нормалізується, за допомогою функції активації (поки що просто уявімо її  $f(x)$ ) і потрапляє в поле  $output$  [10].

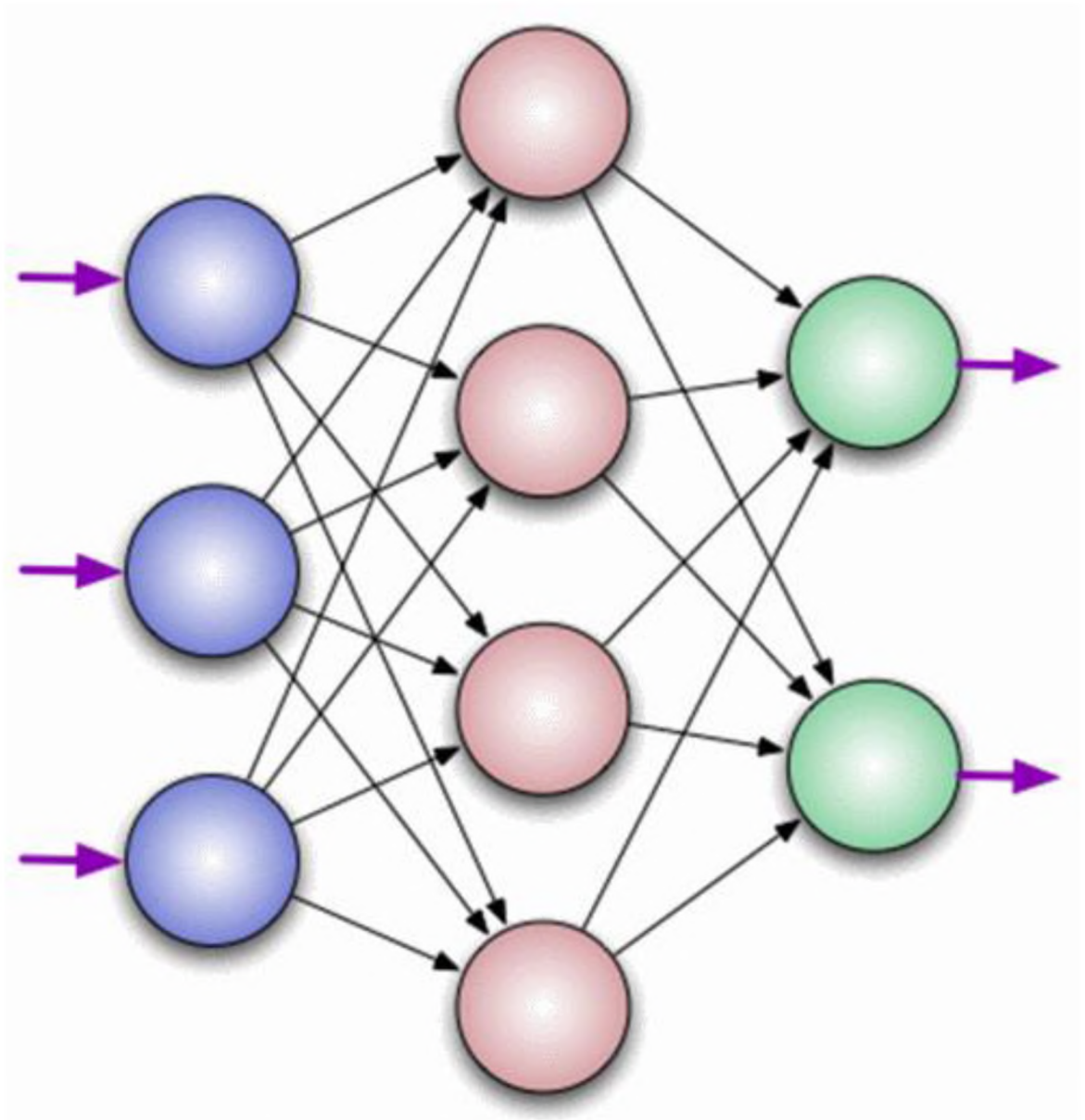


Рисунок 2.3 – Нейронна мережа

Нейрони оперують числами в діапазоні  $[0,1]$  або  $[-1,1]$ . А як же, ви запитаете, тоді обробляти числа, які виходять з даного діапазону? На даному етапі, найпростіша відповідь - це розділити 1 на це число. Цей процес називається нормалізацією, і він дуже часто використовується в нейронних мережах.

Синапс це зв'язок між двома нейронами. У синапсів є 1 параметр - вага. Завдяки йому, вхідна інформація змінюється, коли передається від одного нейрона до іншого. Припустимо, є 3 нейрона, які передають інформацію з наступних підстав. Тоді у нас є 3 ваги, відповідні кожному з цих нейронів як представлено на Рисунку 2.4. У того нейрона, у якого вага буде більше, та



інформація і буде домінуючою в наступному нейроні. Насправді, сукупність ваг нейронної мережі або матриця ваг - це своєрідний мозок всієї системи. Саме завдяки цим ваг, вхідна інформація обробляється і перетворюється в результат.

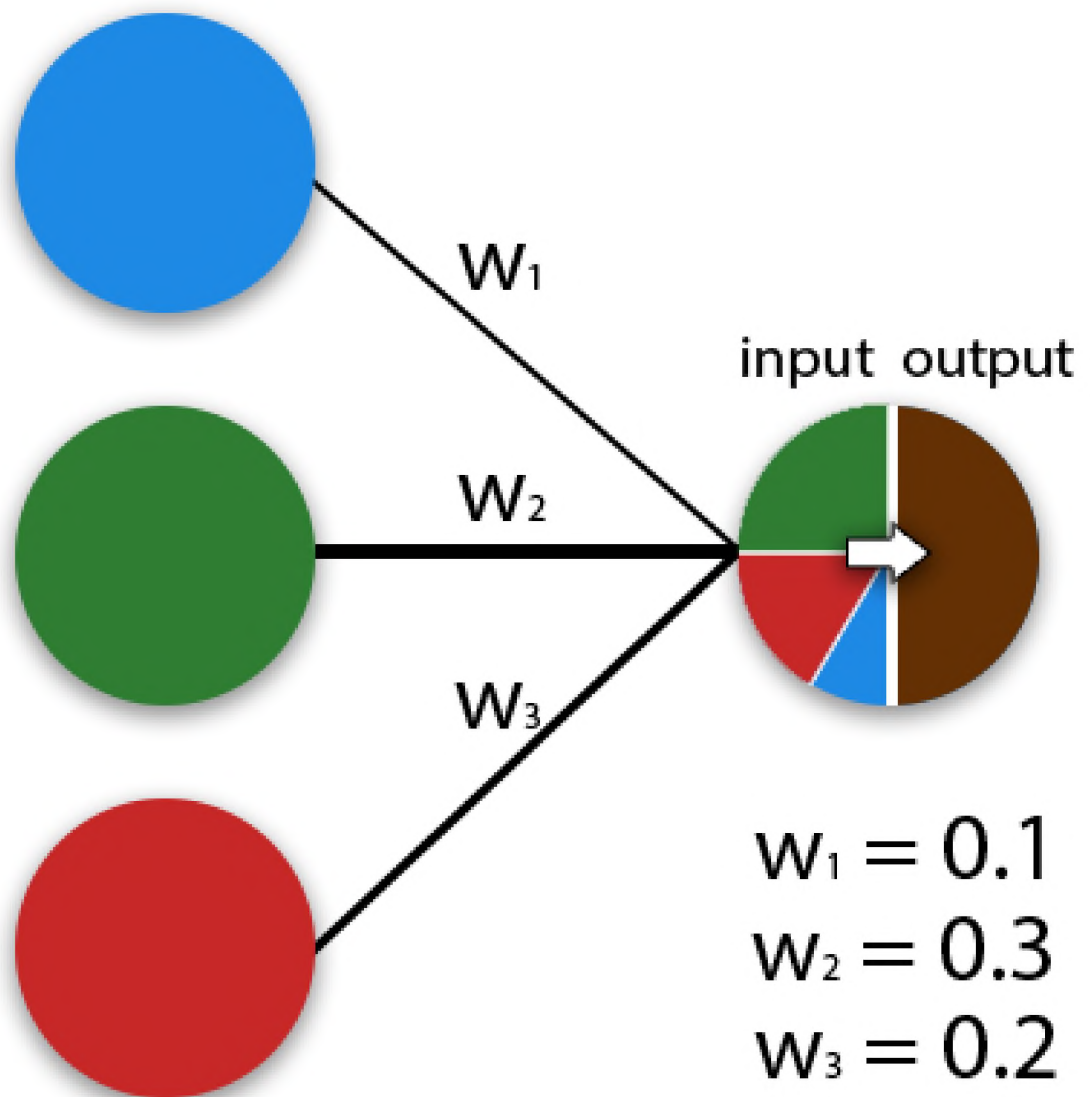
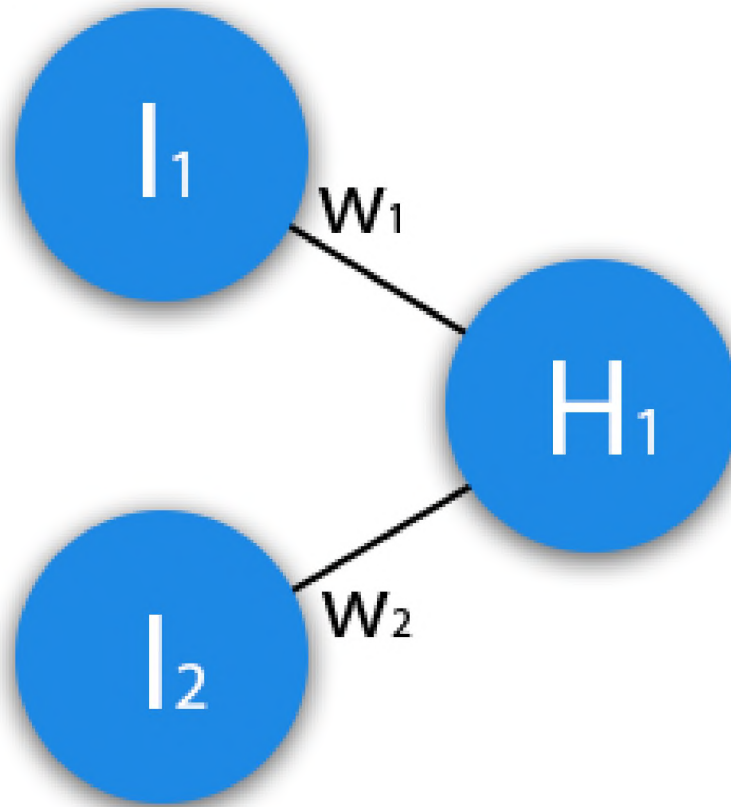


Рисунок 2.4 – Синапс

В прикладі на Рисунок 2.5 зображена частина нейронної мережі, де буквами І позначені вхідні нейрони, буквою Н - прихований нейрон, а буквою w - ваги. З формули видно, що вхідна інформація - це сума всіх вхідних даних, помножених на відповідні їм ваги. Тоді дамо на вхід 1 і 0. Нехай  $w_1 = 0.4$  і  $w_2 = 0.7$  Вхідні дані нейрона Н1 будуть наступними:  $1 * 0.4 + 0 * 0.7 = 0.4$ . Отже, коли у нас є вхідні дані, ми можемо отримати вихідні дані, підставивши вхідний значення в функцію активації. Тепер, коли у нас є



вихідні дані, ми передаємо їх далі. І так, ми повторюємо для всіх верств, поки не дійдемо до вихідного нейрона. Запустивши таку мережу в перший раз ми побачимо, що відповідь далекий від правильно, тому що мережа не натренована. Щоб поліпшити результати ми будемо її тренувати.



$$1) H_{1\text{input}} = (I_1 * w_1) + (I_2 * w_2)$$

$$2) H_{1\text{output}} = f_{\text{activation}}(H_{1\text{input}})$$

Рисунок 2.5 – Приклад роботи нейронної мережі

Функція активації - це спосіб нормалізації вхідних даних. Тобто, якщо на вході у вас буде велика кількість, пропустивши його через функцію активації, ви отримаєте вихід в потрібному вам діапазоні. Функцій активації досить багато тому ми розглянемо найосновніші: Лінійна, сигмоид

(Логістична) і Гіперболічний тангенс. Головні їх відмінності - це діапазон значень.[10]

Лінійна функція представлена на Рисунку 2.6 ця функція майже ніколи не використовується, за винятком випадків, коли потрібно протестувати нейронну мережу або передати значення без перетворень.

$$f(x) = x$$

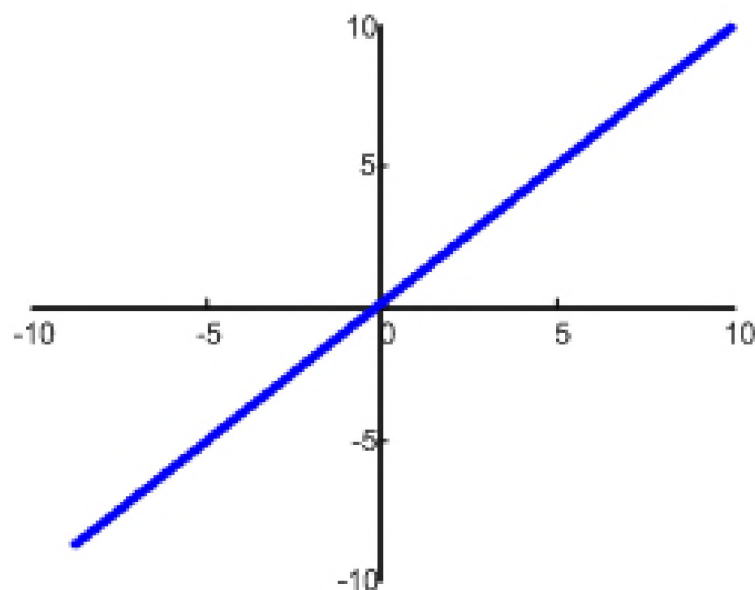


Рисунок 2.6 – Лінійна функція

Сигмоїд представлений на Рисунку 2.7. Це найпоширеніша функція активації, її діапазон значень  $[0,1]$ . Саме на ній показано більшість прикладів в мережі, також її іноді називають логістичною функцією. Відповідно, якщо в вашому випадку присутні негативні значення (наприклад, акції можуть йти не

тільки вгору, але і вниз), то вам знадобитися функція яка захоплює і негативні значення.

$$f(x) = \frac{1}{1 + e^{-x}}$$

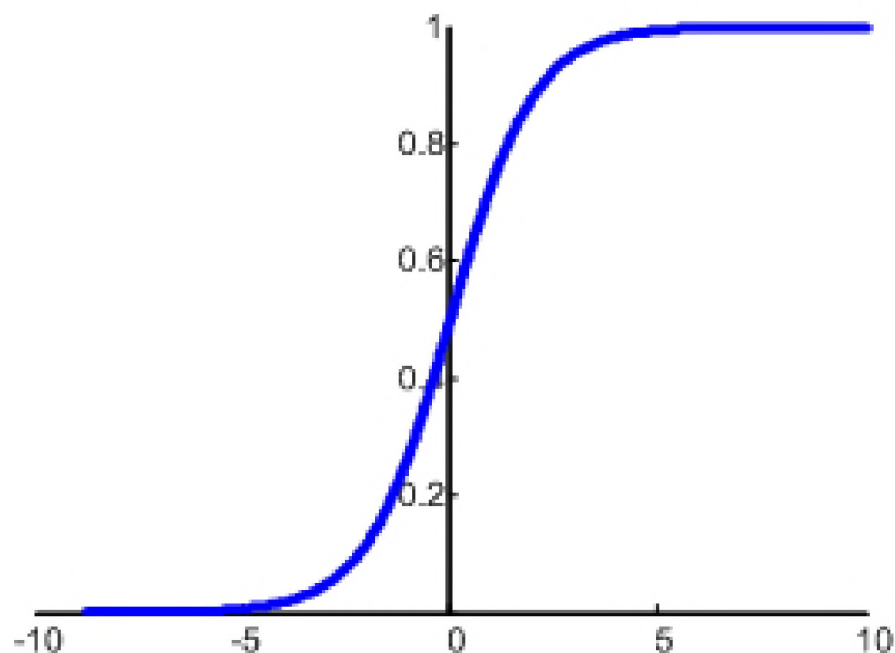


Рисунок 2.7 – Сигмоїд

Гіперболічний тангенс представлений на Рисунок 2.8. Має сенс використовувати гіперболічний тангенс, тільки тоді, коли ваші значення можуть бути і негативними, і позитивними, так як діапазон функції  $[-1,1]$ . Використовувати цю функцію тільки з позитивними значеннями недоцільно так як це значно погіршить результати вашої нейромережі.

$$f(x) = \frac{e^{2x} - 1}{e^{2x} + 1}$$

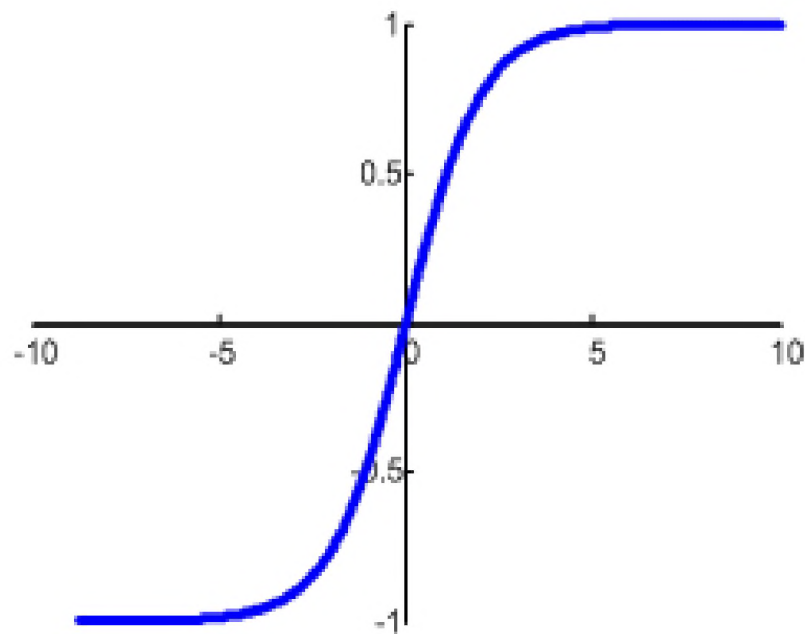


Рисунок 2.8 – Гіперболічний тангенс

Тренувальний сет - це послідовність даних, якими оперує нейронна мережа.

Ітерація це своєрідний лічильник, який збільшується кожного разу, коли нейронна мережа проходить один тренувальний сет. Іншими словами, це загальна кількість тренувальних сетів пройдених нейронною мережею.

Епоха. При ініціалізації нейронної мережі ця величина встановлюється в 0 і має стелю, що задається вручну. Чим більше епоха, тим краще

натренована мережу і відповідно, її результат. Епоха збільшується кожного разу, коли ми проходимо весь набір тренувальних сетів.

Важливо не плутати ітерацію з епохою і розуміти послідовність їх інкремента. спочатку  $n$  раз збільшується ітерація, а потім вже епоха і ніяк не навпаки. Іншими словами, не можна спочатку тренувати нейронну мережу тільки на одному сеті, потім на іншому і так далі. Потрібно тренувати кожен сет один раз за добу. Так, ви зможете уникнути помилок в обчисленнях.

Помилка - це процентна величина, що відображає розбіжність між очікуваним і отриманим відповідями. Помилка формується кожну епоху і повинна йти на спад. Якщо цього не відбувається, значить, ви щось робите не так. Помилку можна обчислити різними шляхами, але ми розглянемо лише три основних способи: Mean Squared Error (далі MSE), Root MSE і Arctan. Стоїть лише враховувати, що кожен метод вважає помилки по-різному. У Arctan, помилка, майже завжди, буде більша, так як він працює за принципом: чим більша різниця, тим більша помилка. У Root MSE буде найменша помилка, тому, найчастіше, використовують MSE, яка зберігає баланс в обчисленні помилки.

Помилка MSE обчислюється за формулою 2.1.

$$\frac{(i_1 - a_1)^2 + (i_2 - a_2)^2 + \dots + (i_n - a_n)^2}{n} \quad (2.1)$$

$i$  – правильна відповідь.

$a$  – отриманий результат

Помилка Root MSE обчислюється за формулою 2.2.

$$\sqrt{\frac{(i_1 - a_1)^2 + (i_2 - a_2)^2 + \dots + (i_n - a_n)^2}{n}} \quad (2.2)$$

Помилка Arctan обчислюється за формулою 2.3.

$$\frac{\arctan^2(i_1 - a_1) + \dots + \arctan^2(i_n - a_n)}{n} \quad (2.3)$$

Принцип підрахунку помилки у всіх випадках однаковий. За кожен сет, ми вважаємо помилку, віднявши від ідеального відповіді, отриманий. Далі, або зводимо в квадрат, або обчислюємо квадратний тангенс з цієї різниці, після чого отримане число ділимо на кількість сетів[10].

Згорткові нейронні мережі. Звучить як дивне поєднання біології та математики з домішкою інформатики, але як би воно не звучало, ці мережі - одні з найвпливовіших інновацій в області комп'ютерного зору. Вперше нейронні мережі привернули загальну увагу в 2012 році, коли Алекс Крижевський завдяки їм виграв конкурс ImageNet.

Сьогодні глибинне навчання лежить в основі послуг багатьох компаній: Facebook використовує нейронні мережі для алгоритмів автоматичного створення тегів, Google - для пошуку серед фотографій користувача, Amazon - для генерації рекомендацій товарів, Pinterest - для персоналізації домашньої сторінки користувача, а Instagram - для пошукової інфраструктури.

Але класичний, і, можливо, найпопулярніший варіант використання мереж це обробка зображень.

Завдання класифікації зображень - це прийом початкового зображення і виведення його класу або групи ймовірних класів, яка найкраще характеризує зображення. Для людей це один з перших навичок, який вони починають освоювати з народження.

Ми опановуємо їм природно, без зусиль, ставши дорослими. Навіть не думаючи, ми можемо швидко і легко розпізнати простір, яке нас оточує, разом з об'єктами. Коли ми бачимо зображення або просто дивимося на те, що відбувається навколо, найчастіше ми можемо відразу охарактеризувати місце дії, дати кожному об'єкту ярлик, і все це відбувається несвідомо, непомітно для розуму. Ці навички швидкого розпізнавання шаблонів, узагальнення вже отриманих знань, і адаптації до різної відображень на фото обстановці не доступні для комп'ютерів[11].

Коли комп'ютер бачить зображення (приймає дані на вхід), він бачить масив пікселів. Залежно від дозволу і розміру зображення, наприклад, розмір

масиву може бути  $32 \times 32 \times 3$  (де 3 - це значення каналів RGB). Щоб було зрозуміліше, давайте уявимо, у нас є кольорове зображення у форматі JPG, і його розмір  $480 \times 480$ . Відповідний масив буде  $480 \times 480 \times 3$ . Кожному з цих чисел присвоюється значення від 0 до 255, яке описує інтенсивність пікселя в цій точці. Ці цифри, залишаючись безглуздими для нас, коли ми визначаємо що на зображенні, є єдиними вступними даними, доступними комп'ютера. Ідея в тому, що ви даєте комп'ютера цю матрицю, а він виводить числа, які описують ймовірність класу зображення.

Ми хочемо, щоб комп'ютер міг розрізняти всі дані йому зображення і розпізнавати унікальні особливості, які роблять собаку собакою, а кішку кішкою. У нас і цей процес відбувається підсвідомо. Коли ми дивимося на зображення собаки, ми можемо віднести його до конкретного класу, якщо у зображення є характерні особливості, які можна ідентифікувати, такі як лапи або чотири ноги. Аналогічним чином комп'ютер може виконувати класифікацію зображень через пошук характеристик базового рівня, наприклад кордонів і викривлень, а потім за допомогою побудови більш абстрактних концепцій через групи згорткових шарів. Це загальний опис того, що роблять згорткові нейронні мережі.

Як ми вже говорили, висновком може бути клас або ймовірність класів, які найкраще описують зображення. Складний момент - розуміння того, що робить кожен з цих шарів. Так що давайте перейдемо до найважливішого[11].

Перший шар в згортковій нейронній мережі завжди згортковий. Як вже говорилося раніше, вступне зображення - це матриця  $32 \times 32 \times 3$  з піксельними значеннями. Найлегше зрозуміти, що таке згортковий шар, якщо уявити його у вигляді ліхтарика, який світить на верхню ліву частину зображення. Припустимо світло, яке випромінює цей ліхтарик, покриває площу  $5 \times 5$ . А тепер давайте уявимо, що ліхтарик рухається по всім областям вступного зображення. У термінах комп'ютерного навчання цей ліхтарик називається фільтром (іноді нейроном або ядром), а області, на які він світить, називаються рецептивних полем. Тобто наш фільтр - це матриця. Зауважте, що глибина у

фільтра повинна бути такою ж, як і глибина вступного зображення, і розміри цього фільтра -  $5 \times 5 \times 3$  фільтр представлений на Рисунок 2.9. Тепер давайте за приклад візьмемо позицію, в якій знаходиться фільтр. Нехай це буде лівий верхній кут. Оскільки фільтр виробляє згортку, тобто пересувається по вступного зображенню, він множить значення фільтра на вихідні значення пікселів зображення. Всі ці множення підсумовуються. І в підсумку виходить одне число. Пам'ятайте, воно просто символізує знаходження фільтра в верхньому лівому кутку зображення. Тепер повторимо цей процес в кожній позиції. (Наступний крок - переміщення фільтра вправо на одиницю, потім ще на одиницю вправо і так далі). Кожна унікальна позиція введеного зображення виробляє число. Після проходження фільтра по всіх позиціях виходить матриця  $28 \times 28 \times 1$ , яку називають функцією активації або картою ознак. Матриця  $28 \times 28$  виходить тому, що є 784 різних позиції, які можуть пройти через фільтр  $5 \times 5$  зображення  $32 \times 32$ . Ці 784 числа перетворюються в матрицю  $28 \times 28$ .

Кожен фільтр можна розглядати як ідентифікатор властивості. Коли я говорю властивість, я маю на увазі прямі кордону, прості кольори і криві. Подумайте про найпростіші характеристиках, які мають всі зображення в загальному. Скажімо, наш перший фільтр  $7 \times 7 \times 3$ , і він буде детектором кривих. У фільтра піксельна структура, в якій чисельні значення вище уздовж області, яка визначає форму кривої[11].

По суті, якщо у ввідному зображенні є форма, в загальних рисах схожа на криву, яку представляє цей фільтр, і всі помножені значення підсумовуються, то результатом буде велике значення.

Останній шар, хоч і знаходиться в кінці, один з важливих. Ми говорили про те, що вміють визначати фільтри першого згорткового шару. Вони виявляють властивості базового рівня, такі як кордони і криві. Як можна собі уявити, щоб припустити який тип об'єкта зображений на картинці, нам потрібна мережа, здатна розпізнавати властивості вищого рівня, як наприклад



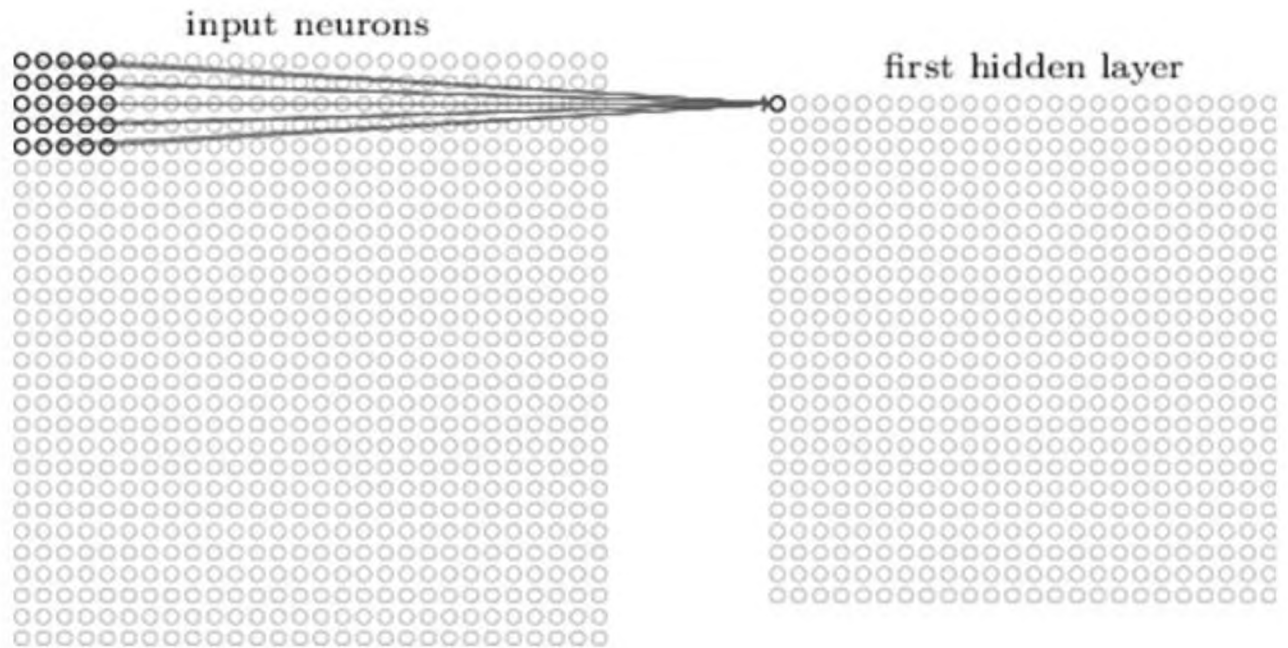


Рисунок 2.9 – зображення фільтра 5x5

руки, лапи або вуха. Так що давайте подумаємо, як виглядає вихідний результат мережі після першого згорткового шару. Його розмір  $28 \times 28 \times 3$  (за умови, що ми використовуємо три фільтри  $5 \times 5 \times 3$ ). Коли картинка проходить через один згортковий шар, вихід першого шару стає вступним значенням 2-го шару. Тепер це трохи складніше візуалізувати. Коли ми говорили про першому шарі, введенням були тільки дані вихідного зображення. Але коли ми перейшли до 2-го шару, вступним значенням для нього стала одна або кілька карт властивостей - результат обробки попереднім шаром. Кожен набір вступних даних описує позиції, де на оригінальному документі зустрічаються певні базові ознаки.

Тепер, коли ми застосовуємо набір фільтрів поверх цього пропускаєте картинку через другий згортковий шар, на виході будуть активовані фільтри, які представляють властивості вищого рівня. Типами цих властивостей можуть бути півкіляця комбінація прямої кордону з вигином або квадратів поєднання декількох прямих ребер. Чим більше згорткових шарів проходить зображення і чим далі воно рухається по мережі, тим складніші характеристики виводяться в картах активації. В кінці мережі можуть бути фільтри, які активуються при наявності рукописного тексту на зображенні, при наявності рожевих об'єктів і так далі.

Тепер, коли ми можемо виявити високорівневі властивості, найкрутіше - це прикріплення повнозв'язну шару в кінці мережі. Цей шар бере ввідні дані і виводить N-просторовий вектор, де N - число класів, з яких програма вибирає потрібний. Наприклад, якщо ви хочете програму з розпізнавання цифр, у N буде значення 10, тому що цифр 10. Кожне число в цьому N-просторовому векторі є ймовірність конкретного класу.

Швидкість навчання - це параметр, який вибирається програмістом. Висока швидкість навчання означає, що в оновленнях ваги робилися більші кроки, тому зразком може знадобитися менше часу, щоб набрати оптимальний набір ваг. Але занадто висока швидкість навчання може привести до дуже великим і недостатньо точним стрибків, які завадять досягненню оптимальних показників[11].

Функція активації ReLU.

Відомо, що нейронні мережі здатні наблизити як завгодно складну функцію, якщо в них досить шарів і функція активації є нелінійною. Функції активації начебто сігмоїдної або тангенціальної є нелінійними, але призводять до проблем з загасанням або збільшенням градієнтів. Однак можна використовувати і набагато простіший варіант - випрямлену лінійну функцію активації (rectified linear unit, ReLU), яка виражається формулою 2.4:

$$f(s) = \max(0, s)$$

(2.4)

Графік функції ReLU відповідно Рисунку 2.10

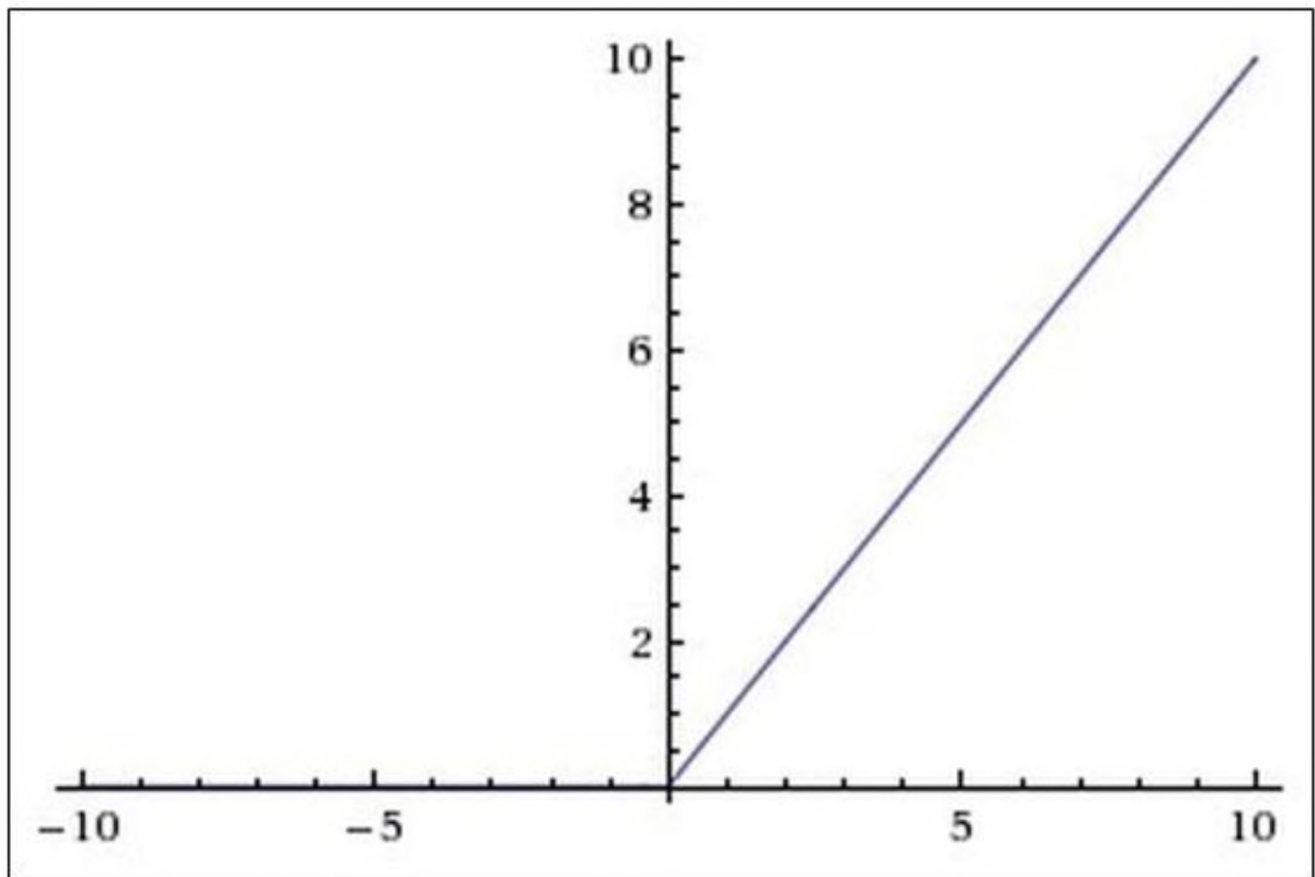


Рисунок 2.10 – Графік функції ReLU

Переваги використання ReLU.

Її похідна дорівнює або одиниці, або нулю, і тому не може статися розростання або загасання градієнтів, тому що помноживши одиницю на дельту помилки ми отримаємо дельту помилки, якщо ж ми б використовували іншу функцію, наприклад, гіперболічний тангенс, то дельта помилки могла, або зменшитися, або зрости, або залишитися такою ж, тобто, похідна гіперболічного тангенса повертає число з різним знаком і величиною, що можна сильно вплинути на загасання або розростання градієнта. Більш того, використання даної функції приводить до проріджування ваг.

Обчислення сигмоїд і гіперболічного тангенса вимагає виконання ресурсномістких операцій, таких як спорудження до рівня, в той час як ReLU може бути реалізований за допомогою простого порогового перетворення матриці активацій в нулі.

Відсікає непотрібні деталі в каналі при негативному виході.

З недоліків можна відзначити, що ReLU не завжди достатньо надійна і в процесі навчання може виходити з ладу. Наприклад, великий градієнт, що проходить через ReLU, може привести до такого оновлення ваг, що даний нейрон ніколи більше не активується. Якщо це станеться, то, починаючи з даного моменту, градієнт, що проходить через цей нейрон, завжди буде дорівнює нулю. Відповідно, даний нейрон буде необоротно виведений з ладу. Наприклад, при дуже великій швидкості навчання (learning rate), може виявитися, що до 40% ReLU «мертві» (тобто, ніколи не активуються). Ця проблема вирішується за допомогою вибору належної швидкості навчання[12].

## 2.5 Часові ряди

Часовий ряд - зібраний в різні моменти часу статистичний матеріал про значення будь-яких параметрів (в найпростішому випадку одного) досліджуваного процесу. Кожна одиниця статистичного матеріалу називається виміром або відліком, також допустимо називати його рівнем на вказаний з ним момент часу. У тимчасовому ряді для кожного відліку має бути зазначено час вимірювання або номер вимірювання по порядку. Часовий ряд істотно відрізняється від простої вибірки даних, так як при аналізі враховується взаємозв'язок вимірювань з часом, а не тільки статистичне різноманітність і статистичні характеристики вибірки.

Часові ряди дуже часто представляють за допомогою лінійних діаграм. Часові ряди використовуються в статистиці, обробці сигналів, розпізнаванні образів, економетриці, фінансовій математиці, прогнозуванні погоди, розумному транспорті та передбаченні траєкторій, передбаченні землетрусів, автоматичному керуванні, астрономії, технологіях зв'язку, а також значною мірою в будь-якій області прикладної науки та інженерії, яка включає часові вимірювання.

На Рисунку 2.11 представлений часовий ряд з лінійним трендом.

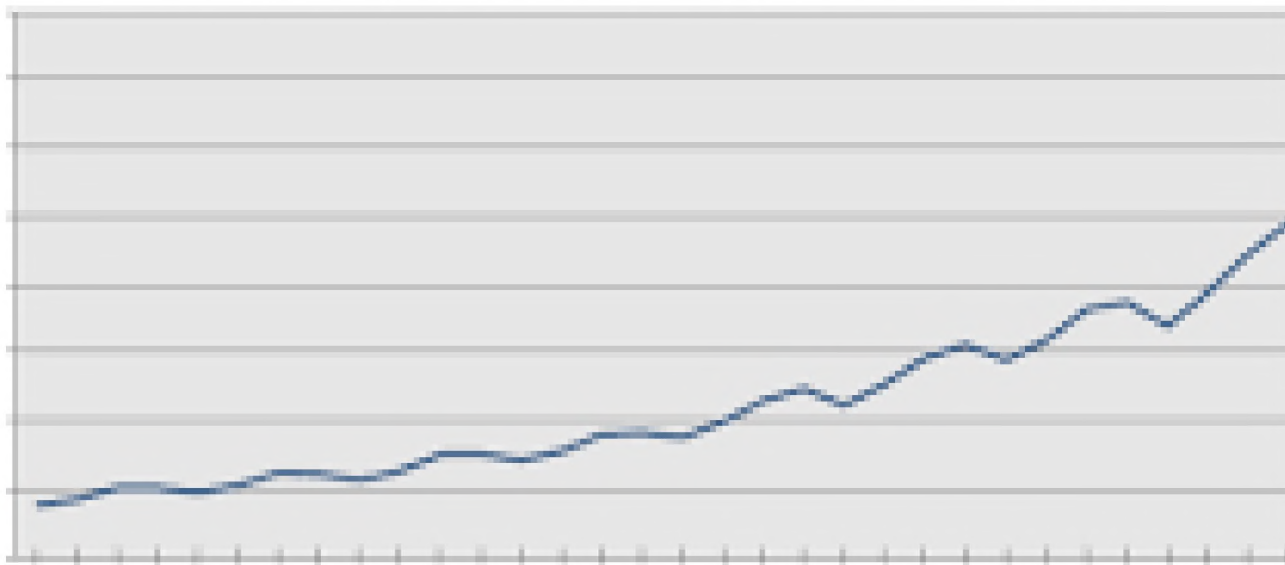


Рисунок 2.11 – Часовий ряд з лінійним трендом

На Рисунку 2.12 представлений часовий ряд з випадковими коливаннями

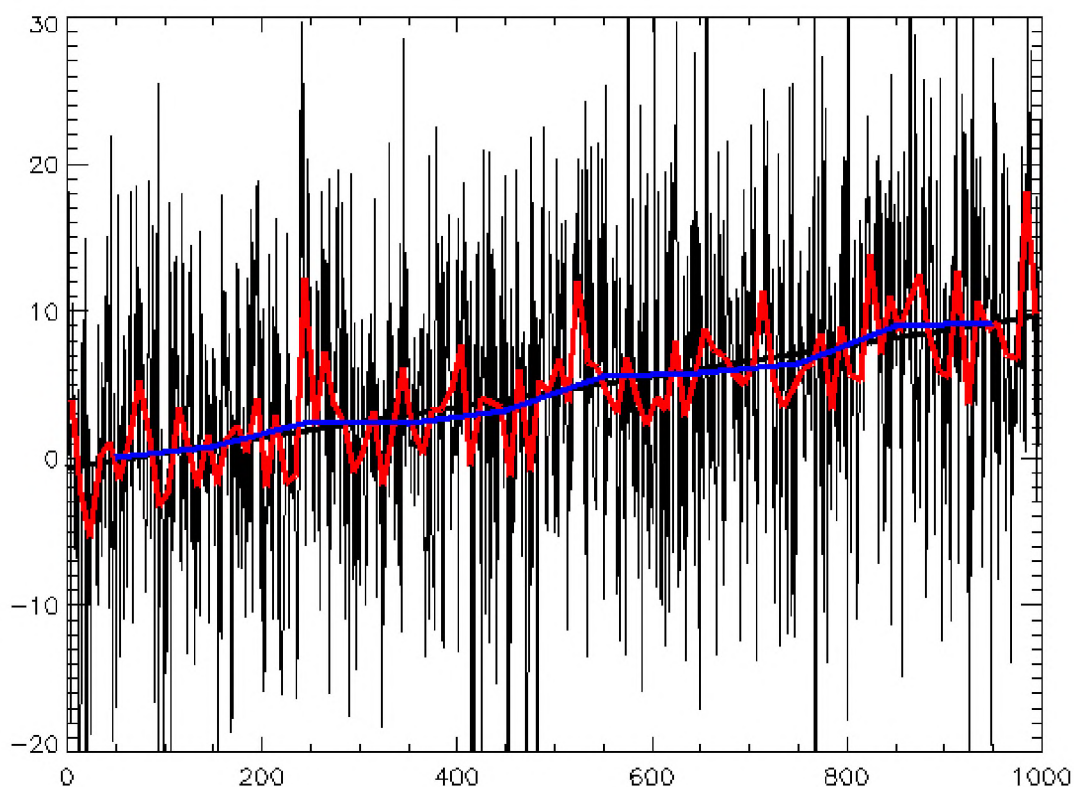


Рисунок 2.12 – графік коливань

Аналіз часових рядів — сукупність математико-статистичних методів аналізу, призначених для виявлення структури часових рядів і для їх прогнозування. Сюди належать, зокрема, методи регресійного аналізу.

Виявлення структури часового ряду необхідно для того, щоб побудувати математичну модель того явища, яке є джерелом аналізованого часового ряду. Прогноз майбутніх значень часового ряду використовується для ефективного прийняття рішень.

Існують дві основні мети аналізу часових рядів:

1. Визначення природи ряду
2. Прогнозування

Обидві ці цілі вимагають, щоб модель ряду була ідентифікована і, більш-менш, формально описана. Як тільки модель визначена, ви можете з її допомогою інтерпретувати представлені дані (наприклад, використовувати у вашій теорії для розуміння сезонного зміни цін на товари, якщо займаєтеся економікою). Не звертаючи уваги на глибину розуміння і справедливості теорії, ви можете екстраполювати потім ряд на основі знайденої моделі, тобто передбачити його майбутні значення.

Часові ряди досліджуються з різними цілями. В одному ряді випадках буває достатньо отримати опис характерних особливостей ряду, а в іншому ряді випадків потрібне не тільки передбачати майбутні значення часового ряду, а й управляти його поведінкою. Метод аналізу часового ряду визначається, з одного боку, цілями аналізу, а з іншого боку, ймовірнісною природою формування його значень[13].

Найпоширеніші методи аналізу часових рядів:

1. Спектральний аналіз — дозволяє знаходити періодичні складові часового ряду
2. Кореляційний аналіз — дозволяє знаходити суттєві періодичні залежності і відповідні їм затримки (лаги) як всередині одного ряду (автокореляція), так і між кількома рядами.
3. Моделі авторегресії і ковзного середнього — моделі орієнтовані на опис процесів, що виявляють однорідні коливання, порушувани випадковими впливами. Дозволяють передбачати майбутні значення ряду.

4. Багатоканальні моделі авторегресії і ковзного середнього — моделі застосовуються в тих випадках, коли є кілька корельованих між собою часових рядів. У них є коливання, порушені однією причиною. Дозволяють передбачати майбутні значення ряду.
5. Сезонна модель Бокса-Дженкінса — застосовується, коли часовий ряд містить явно виражений лінійний тренд і сезонні складові. Дозволяє передбачати майбутні значення ряду. Модель була запропонована у зв'язку з аналізом авіаперевезень.
6. Прогноз експоненціально зваженим ковзаючим середнім — найпростіша модель прогнозування часового ряду. Застосовна в багатьох випадках. У тому числі, охоплює модель ціноутворення на основі випадкових блукань.

Зазвичай при практичному аналізі часових рядів послідовно проходять такі етапи:

1. Графічне подання і опис поведінки часового ряду.
2. Виділення та видалення закономірних складових часового ряду, що залежать від часу: тренда, сезонних і циклічних складових.
3. Виділення та видалення низько- або високочастотних складових процесу (фільтрація).
4. Дослідження випадкової складової часового ряду, що залишилася після видалення перерахованих вище складових.
5. Побудова математичної моделі для опису випадкової складової і перевірка її адекватності.
6. Прогнозування майбутнього розвитку процесу, представленого часовим рядом.

Як і більшість інших видів аналізу, аналіз часових рядів передбачає, що дані містять систематичну складову зазвичай включає кілька компонент і випадковий шум, який ускладнює виявлення регулярних компонент. Більшість методів дослідження часових рядів включає різні способи фільтрації шуму, що дозволяють побачити регулярну складову більш чітко.

Часовий ряд — це послідовність впорядкованих у часі числових показників, що характеризують рівень стану і зміни досліджуваного явища.

Всякий часовий ряд включає два обов'язкових елемента: по-перше, час і, по-друге, конкретне значення показника, або рівень ряду[13].

Часові ряди розрізняються за такими ознаками: за часом — моментні та інтервальні. Інтервальний ряд — послідовність, в якій рівень явища відносять до результату, накопиченому або знову виробленому за певний інтервал часу. Такі, наприклад, такі ряди показників обсягу продукції підприємства по місяцях року, кількості відпрацьованих людиною днів по окремих періодах. Якщо ж рівень ряду характеризує досліджуване явище в конкретний момент часу, то сукупність рівнів формує моментний ряд.

Найважливішою умовою правильного формування часових рядів є порівнянність рівнів, що утворюють ряд. Рівні ряду, що підлягають вивченню, повинні бути однорідні за економічним змістом і враховувати сутність досліджуваного явища і його мету.

Статистичні дані, представлені у вигляді часових рядів, повинні бути порівняні по території, колу охоплюваних об'єктів, одиницях вимірювання, моменту реєстрації, методикою розрахунку, цінами, достовірності. Неспівмірність за територією виникає в результаті змін кордонів країн, регіонів, господарств тощо. Для приведення даних до порівнянного вигляду проводиться перерахунок колишніх даних з урахуванням нових кордонів.

Цілісність охоплення різних частин явища — найважливіша умова порівнянності рівнів ряду. Вимога однакової цілісності охоплення різних частин досліджуваного об'єкта означає, що рівні ряду за окремі періоди повинні характеризувати різноманітність заходів того чи іншого явища по одному і тому ж колу, які входять до складу його частин. Наприклад, при характеристиці динаміки врожайності овочевих культур у регіоні по роках не можна в одні роки враховувати тільки сільськогосподарські підприємства, а в інші — всі категорії господарств, у тому числі присадибні ділянки сільських жителів і сади, городи городян[13].



При визначенні порівнюваних рівнів ряду необхідно використовувати єдину методику їх розрахунку. Особливо часто ця проблема виникає при міжнародних порівняннях.

При аналізі показників у вартісному вираженні слід враховувати, що з плином часу відбувається безперервна зміна цін. Причин цього процесу безліч — інфляція, зростання витрат, ринкові умови (попит і пропозиція) і так далі. У цьому зв'язку при характеристиці вартісних показників обсягів продукції в часі має бути усунуто вплив трансформаційних змін цін. Для вирішення цього завдання кількість продукції, виробленої в різні періоди, оцінюють в цінах одного періоду, які називають фіксованими або в визначених в статистичних органах — порівнянними цінами[13].

Широке використання в статистичних дослідженнях вибіркового методу вимагає враховувати достовірність кількісних і якісних характеристик досліджуваних явищ в динаміці. Різна репрезентативність вибірки за періодами внесе суттєві похибки в величини рівнів ряду.

Однією з умов порівнянності рівнів інтервального ряду, крім рівності періодів, за які наводять дані, є однорідність етапів, у межах яких показник підпорядковується одному закону розвитку. У цих випадках проводять періодизацію часових рядів, типологічну угруповання в часі. Всі вищеназвані обставини слід враховувати при підготовці інформації для аналізу змін явищ у часі[13].

Зазвичай, метою прикладного статистичного аналізу часових рядів є побудова моделі ряду, за допомогою якої можна пояснити поведінку ряду і здійснити прогноз на майбутні періоди.

Аналіз часового ряду починається з побудови і вивчення його графіка. Якщо нестационарність часового ряду очевидна, то спочатку необхідно виокремити його нестационарну складову. Процес виокремлення тренду та інших компонент ряду, що призводять до порушення стаціонарності, може проходити в декілька етапів. На кожному з них розглядається ряд залишків, отриманий у результаті вирахування з вихідного ряду підібраної моделі

тренду, або результат різницевих і інших перетворень ряду. Крім графіків, ознаками нестационарності часового ряду можуть служити автокореляційна функція, що прямує не до нуля і наявність яскраво виражених піків на низьких частотах у періодограмі. За допомогою автокореляційної функції досліджують також внутрішні зв'язки між елементами часових рядів.

У вибіркових дослідженнях найпростіші числові характеристики описової статистики звичайно дають достатньо інформативне уявлення про вибірку. Графічні методи зображення й аналізу вибірок при цьому грають лише допоміжну роль, дозволяючи краще зрозуміти локалізацію і концентрацію даних, їхній закон розподілу.

Роль графічних методів при аналізі часових рядів цілком інша. Табличне представлення часового ряду й описових статистик частіше за все не дозволяє зрозуміти характер процесу, у той час як за графіком часового ряду можна зробити досить багато висновків. Надалі вони можуть бути перевірені й уточнені за допомогою розрахунків.

Людське око досить упевнено визначає за графіком часового ряду:

1. Наявність тренду і його характер.
2. Наявність сезонних і циклічних компонент.
3. Ступінь повільності або переривчастості змін послідовних значень ряду після усунення тренду.

Так графічний аналіз ряду звичайно задає напрямок його подальшого аналізу.

Останнім етапом аналізу часового ряду може бути прогнозування його майбутніх або відновлення пропущених значень і визначення точності цього прогнозу на базі підібраної моделі. Добре підібрати математичну модель вдається не для всякого часового ряду. Нерідко буває і так, що для опису підходять відразу декілька моделей. Неоднозначність вибору моделі може спостерігатися як на етапі виділення детермінованого компонента ряду, так і при виборі структури ряду залишків. Тому досить часто розробляють декілька прогнозів, зроблених за допомогою різних моделей[13].

## **Висновки до розділу 2**

Використовуючи Twitter API можливо зібрати необхідні данні по вибраним користувачам з соціальної мережі Twitter.

Для аналізу зібраних даних можливо використати нейронні мережі для задачі класифікації даних на три класи.

Та представити отримані результати у вигляді часового ряду.

## 3 ПРАКТИЧНА ЧАСТИНА

### 3.1 Збір даних

Для збору даних була вибрана соціальна мережа Twitter. Була створена прикладна програма в Twitter завдяки якій було отримано доступ до даних соціальної мережі. Для підключення до Twitter API мені були надані ключ споживача, та секретний ключ споживача. Та маркер доступу та секретний маркер доступу.

Для завантаження даних було обрано п'ять груп з новинами України:

1. 112 News Feed
2. KanalUkraine\_en
3. News on Ukraine
4. Ua News online
5. Unian Info

Була написана програма на мові програмування Python для завантаження даних з Twitter взаємодіючи з Twitter API.

```
user = sys.argv[1]
client = get_twitter_client()
fname = "user_timeline_().jsonl".format(user)
with open(fname, 'w') as f:
    for page in Cursor(client.user_timeline,
                        screen_name=user,
                        count=200).pages(16):
        for status in page:
            f.write(json.dumps(status._json)+"\n")
```

Приклад одного завантаженого твіту.

```
{
  "created_at": "Tue Nov 20 17:01:50 +0000 2018",
  "id": 1064926806598406144, "id_str": "1064926806598406144",
```

```

"text": "#Russia loses battle for the #OrthodoxChurch with
#Ukraine\nhttps://t.co/Ly5H5TrmqO",
"truncated": false,
"entities": {
"hashtags": [{ "text": "Russia", "indices": [0, 7]},
{ "text": "OrthodoxChurch", "indices": [29, 44]},
{ "text": "Ukraine", "indices": [50, 58]}],
"symbols": [],
"user_mentions": [],
"urls": [{ "url": "https://t.co/Ly5H5TrmqO",
"expanded_url": "https://112.international/opinion/russia-loses-battle-for-
the-orthodox-church-with-ukraine-34351.html",
"display_url": "112.international/opinion/russia\u2026", "indices": [59,
82]}]},
"source": "<a href=\"http://twitter.com\" rel=\"nofollow\">Twitter Web
Client</a>",
"in_reply_to_status_id": null,
"in_reply_to_status_id_str": null,
"in_reply_to_user_id": null,
"in_reply_to_user_id_str": null,
"in_reply_to_screen_name": null,
"user": { "id": 3390749657,
"id_str": "3390749657",
"name": "112 News",
"screen_name": "112NewsFeed",
"location": "Kyiv",
"description": "News from Ukraine. Fast. Reliable. Exclusive. Everything
you need to know about Ukraine.",
"url": "https://t.co/8QcoiV7kqb",

```

```

"entities": {"url": {"urls": [{"url": "https://t.co/8QcoiV7kqb", "expanded_url":
"https://112.international/", "display_url": "112.international", "indices": [0,
23]}]},
  "description": {"urls": []}},
  "protected": false,
  "followers_count": 1966,
  "friends_count": 312,
  "listed_count": 175,
"created_at": "Fri Jul 24 09:57:31 +0000 2015", "favourites_count": 173,
  "utc_offset": null,
  "time_zone": null,
  "geo_enabled": true,
  "verified": false,
  "statuses_count": 31432,
  "lang": "en",
  "contributors_enabled": false,
  "is_translator": false,
  "is_translation_enabled": false,
"profile_background_color": "C0DEED", "profile_background_image_url":
"http://abs.twimg.com/images/themes/theme1/bg.png",
"profile_background_image_url_https":
"https://abs.twimg.com/images/themes/theme1/bg.png",
"profile_background_tile": false,
"profile_image_url":
"http://pbs.twimg.com/profile_images/633703856120180737/QSe6R1kh_normal.p
ng",
"profile_image_url_https":
"https://pbs.twimg.com/profile_images/633703856120180737/QSe6R1kh_normal.
png",

```

```

"profile_banner_url":
"https://pbs.twimg.com/profile_banners/3390749657/1538654064",
"profile_link_color": "1DA1F2", "profile_sidebar_border_color": "C0DEED",
"profile_sidebar_fill_color": "DDEEF6",
"profile_text_color": "333333", "profile_use_background_image": true,
"has_extended_profile": false,
"default_profile": true,
"default_profile_image": false,
"following": true,
"follow_request_sent": false,
"notifications": false, "translator_type": "none"},
  "geo": null,
  "coordinates": null,
  "place": null,
  "contributors": null,
  "is_quote_status": false,
  "retweet_count": 0,
  "favorite_count": 0,
  "favorited": false,
  "retweeted": false,
  "possibly_sensitive": false,
  "lang": "en"}

```

З кожного твіту необхідні дані знаходяться в полях "created\_at" та "text". В результаті отримаємо набір даних, який складається з дати створення твіту та його тексту.

### 3.2 Аналіз даних

Для проведення класифікації твітів за трьома типами емоцій була створена нейронна мережа за допомогою відкритої бібліотеки від компанії Google TensorFlow.

Нейронна мережа має два прихованих шари та наступні параметри.

```
# Network Parameters
n_hidden_1 = 10  # 1st layer number of features
n_hidden_2 = 10  # 2nd layer number of features
n_input = total_words # Words in vocab
n_classes = 3

input_tensor = tf.placeholder(tf.float32,[None, n_input],name="input")
output_tensor = tf.placeholder(tf.float32,[None, n_classes],name="output")

def multilayer_perceptron(input_tensor, weights, biases):
    layer_1_multiplication = tf.matmul(input_tensor, weights['h1'])
    layer_1_addition = tf.add(layer_1_multiplication, biases['b1'])
    layer_1 = tf.nn.relu(layer_1_addition)

    # Hidden layer with RELU activation
    layer_2_multiplication = tf.matmul(layer_1, weights['h2'])
    layer_2_addition = tf.add(layer_2_multiplication, biases['b2'])
    layer_2 = tf.nn.relu(layer_2_addition)

    # Output layer
    out_layer_multiplication = tf.matmul(layer_2, weights['out'])
    out_layer_addition = out_layer_multiplication + biases['out']
    return out_layer_addition
```



Ваги і зміщення зберігаються в змінних `tf.Variable`, які містять стан в графі між викликами. Ваги та зміщення отримуються через нормальний розподіл.

```
weights = {  
    'h1': tf.Variable(tf.random_normal([n_input, n_hidden_1])),  
    'h2': tf.Variable(tf.random_normal([n_hidden_1, n_hidden_2])),  
    'out': tf.Variable(tf.random_normal([n_hidden_2, n_classes]))  
}  
biases = {  
    'b1': tf.Variable(tf.random_normal([n_hidden_1])),  
    'b2': tf.Variable(tf.random_normal([n_hidden_2])),  
    'out': tf.Variable(tf.random_normal([n_classes]))  
}
```

Для навчання моделі було класифіковано вручну 126 твітів. Дані твіти були розбиті на дві частини для навчання і тестування моделі. Отримана точність моделі дорівнює 65%.

### 3.3 Результати

#### 3.3.1 Результат групи 112 News

Результати проведення класифікації твітів групи 112 News представленні у Таблиці 3.1. А саме кількість твітів з емоцією надії, страху та приниження та загальною кількістю твітів. Дані розбиті на 10 тижнів починаючи з 12.09.2018 по 20.11.2018.

Таблиця 3.1 – Класифіковані дані групи 112 News

112 News				
	Надія	Страх	Приниження	Всього
12.09.2018-18.09.2018	63	74	5	142
19.09.2018-25.09.2018	63	67	10	140
26.09.2018-2.10.2018	61	62	13	136
3.10.2018-9.10.2018	67	51	12	130
10.10.2018-16.10.2018	31	34	4	69
17.10.2018-23.10.2018	59	68	12	139
24.10.2018-30.10.2018	70	56	11	137
31.10.2018-6.11.2018	64	67	6	137
7.11.2018-13.11.2018	62	67	7	136
14.11.2018-20.11.2018	28	36	3	67

За отриманими результатами було побудовано часовий ряд який представлений на Рисунку 3.1.

Були визначенні функції ліній тренду у поліноміальній формі 2-го ступеню.

$$y = -0,1926x^2 + 2,0363x + 42,047 \text{ – для емоції надія}$$

$$y = 0,3853x^2 - 3,9616x + 54,532 \text{ – для емоції страх}$$

$$y = -0,1927x^2 + 1,9253x + 3,4204 \text{ – для емоції приниження}$$

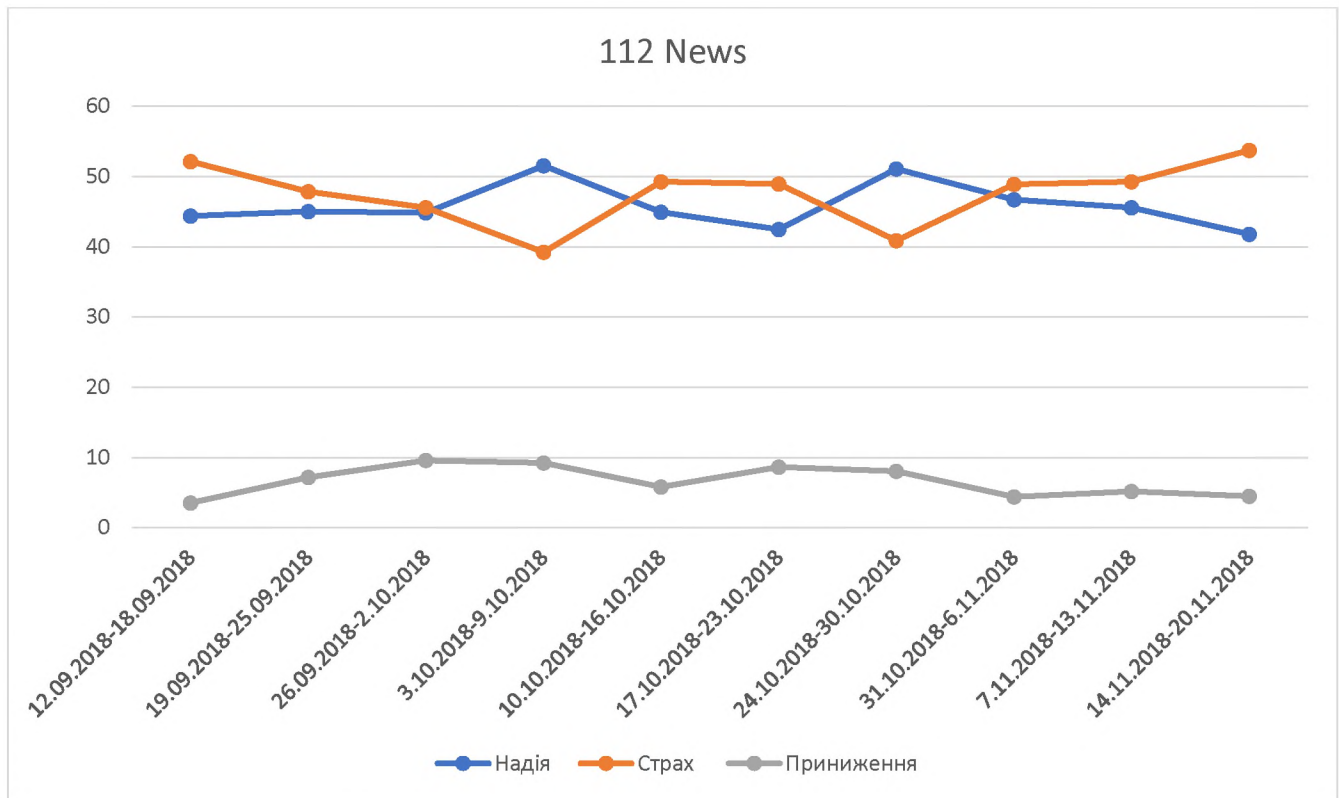


Рисунок 3.1 – Часовий ряд за даними групи 112 News

### 3.3.2 Результат групи UnianInfo

Результати проведення класифікації твітів групи UnianInfo представлені у Таблиці 3.2. А саме кількість твітів з емоцією надії, страху та приниження та загальною кількістю твітів. Дані розбиті на 10 тижнів починаючи з 12.09.2018 по 20.11.2018.

Таблиця 3.2 - Класифіковані дані групи UnianInfo

UnianInfo				
	Надія	Страх	Приниження	Всього
12.09.2018-18.09.2018	69	43	10	122
19.09.2018-25.09.2018	58	55	11	124
26.09.2018-2.10.2018	50	65	7	122
3.10.2018-9.10.2018	62	47	15	124
10.10.2018-16.10.2018	70	31	23	124
17.10.2018-23.10.2018	55	45	25	125
24.10.2018-30.10.2018	58	50	19	127
31.10.2018-6.11.2018	84	83	22	189
7.11.2018-13.11.2018	73	83	27	183
14.11.2018-20.11.2018	60	43	21	124

За отриманими результатами було побудовано часовий ряд який представлений на Рисунку 3.2.

Були визначенні функції ліній тренду у поліноміальній формі 2-го ступеню.

$$y = 0,1089x^2 - 1,9851x + 54,041 \text{ – для емоції надії}$$

$$y = 0,0983x^2 - 1,2601x + 42,657 \text{ – для емоції страху}$$

$$y = -0,2072x^2 + 3,2452x + 3,3029 \text{ – для емоції приниження}$$

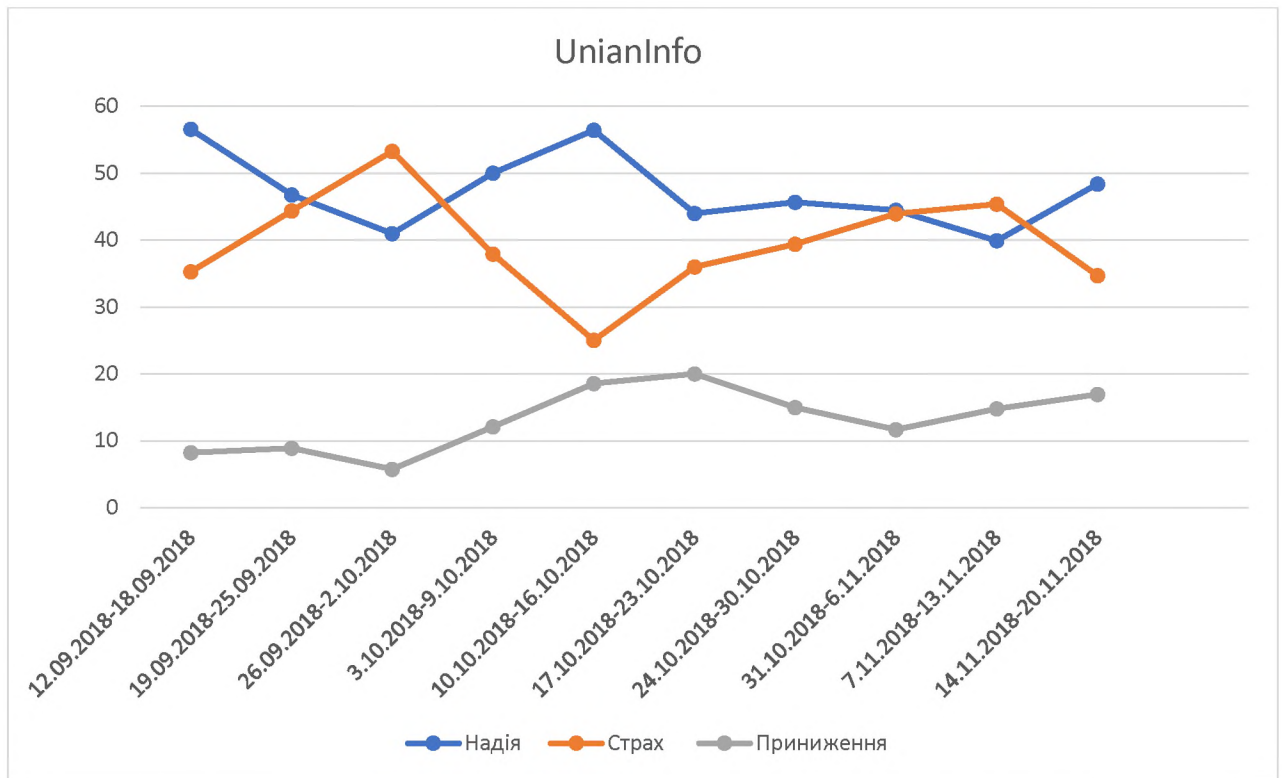


Рисунок 3.2 – Часовий ряд за даними групи UnianInfo

### 3.3.3 Результат групи UaNews

Результати проведення класифікації твітів групи UaNews представлені у Таблиці 3.3. А саме кількість твітів з емоцією надії, страху та приниження та загальною кількістю твітів. Дані розбиті на 10 тижнів починаючи з 12.09.2018 по 20.11.2018.

Таблиця 3.3 – Класифіковані дані групи UaNews

UaNews				
	Надія	Страх	Приниження	Всього
12.09.2018-18.09.2018	11	9	2	22
19.09.2018-25.09.2018	5	6	2	13
26.09.2018-2.10.2018	6	6	0	12
3.10.2018-9.10.2018	17	13	4	34
10.10.2018-16.10.2018	21	10	3	34
17.10.2018-23.10.2018	19	11	3	33
24.10.2018-30.10.2018	19	9	4	32
31.10.2018-6.11.2018	11	10	2	23
7.11.2018-13.11.2018	9	7	2	18
14.11.2018-20.11.2018	10	6	1	17

За отриманими результатами було побудовано часовий ряд який представлений на Рисунку 3.3.

Були визначенні функції ліній тренду у поліноміальній формі 2-го ступеню.

$$y = -0,2782x^2 + 4,1103x + 40,487 \text{ – для емоції надія}$$

$$y = 0,3173x^2 - 4,4624x + 50,71 \text{ – для емоції страху}$$

$$y = -0,0391x^2 + 0,352x + 8,8027 \text{ – для емоції приниження}$$

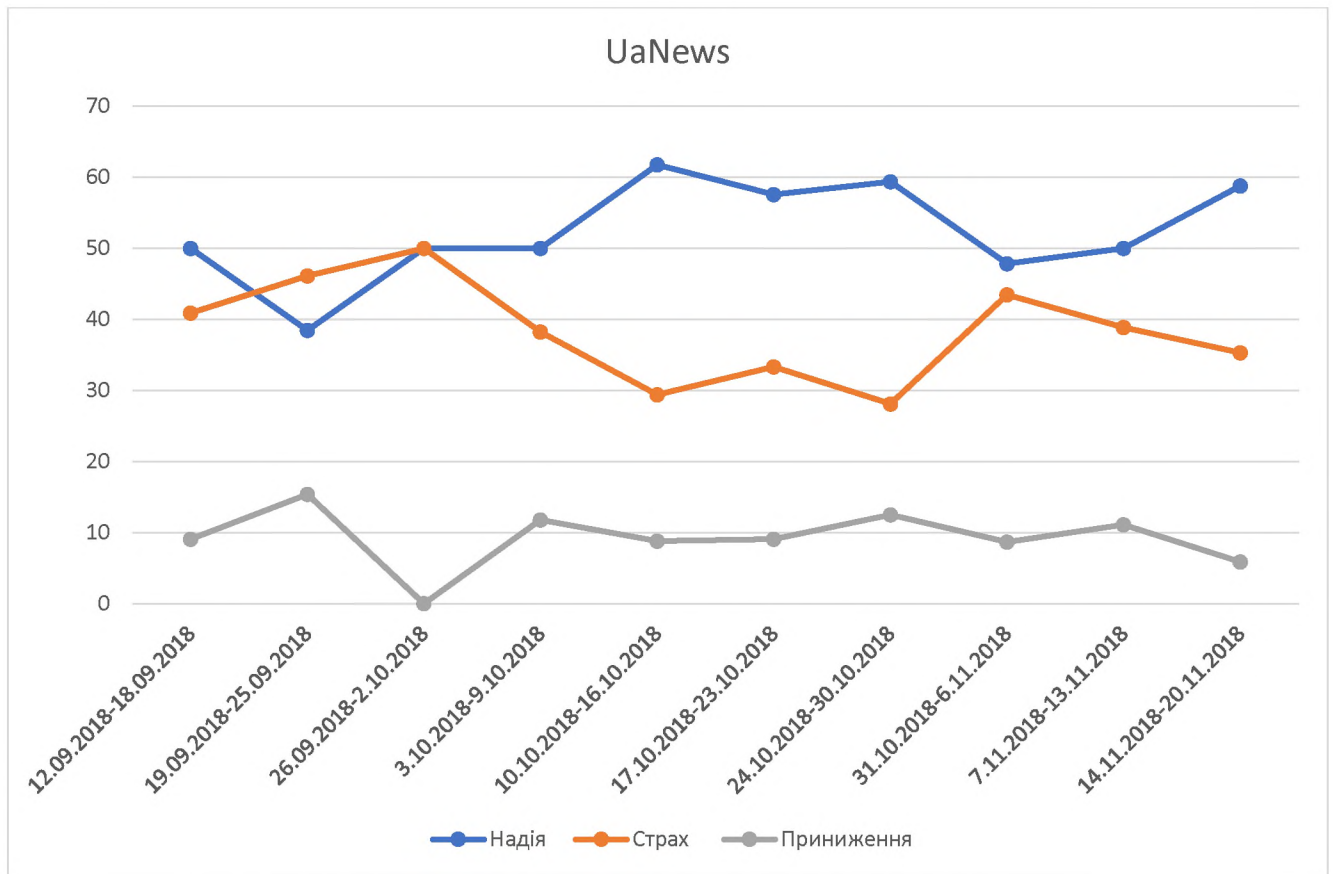


Рисунок 3.3 – Часовий ряд за даними групи UaNews

### 3.3.4 Результат групи kanalukraine

Результати проведення класифікації твітів групи kanalukraine представленні у Таблиці 3.4. А саме кількість твітів з емоцією надії, страху та приниження та загальною кількістю твітів. Дані розбиті на 10 тижнів починаючи з 12.09.2018 по 20.11.2018.

Таблиця 3.4 – Класифіковані дані групи kanalukraine

kanalukraine				
	Надія	Страх	Приниження	Всього
12.09.2018-18.09.2018	10	4	2	16
19.09.2018-25.09.2018	4	5	2	11
26.09.2018-2.10.2018	2	2	0	4
3.10.2018-9.10.2018	9	5	1	15
10.10.2018-16.10.2018	4	2	2	8
17.10.2018-23.10.2018	6	3	1	10
24.10.2018-30.10.2018	3	3	2	8
31.10.2018-6.11.2018	5	2	3	10
7.11.2018-13.11.2018	4	3	1	8
14.11.2018-20.11.2018	6	2	0	8

За отриманими результатами було побудовано часовий ряд який представлений на Рисунку 3.4.

Були визначенні функції ліній тренду у поліноміальній формі 2-го ступеню.

$$y = 0,6259x^2 - 5,9726x + 61,89 \text{ – для емоції надії}$$

$$y = -0,1386x^2 + 0,3841x + 36,102 \text{ – для емоції страху}$$

$$y = -0,4873x^2 + 5,5885x + 2,0076 \text{ – для емоції приниження}$$



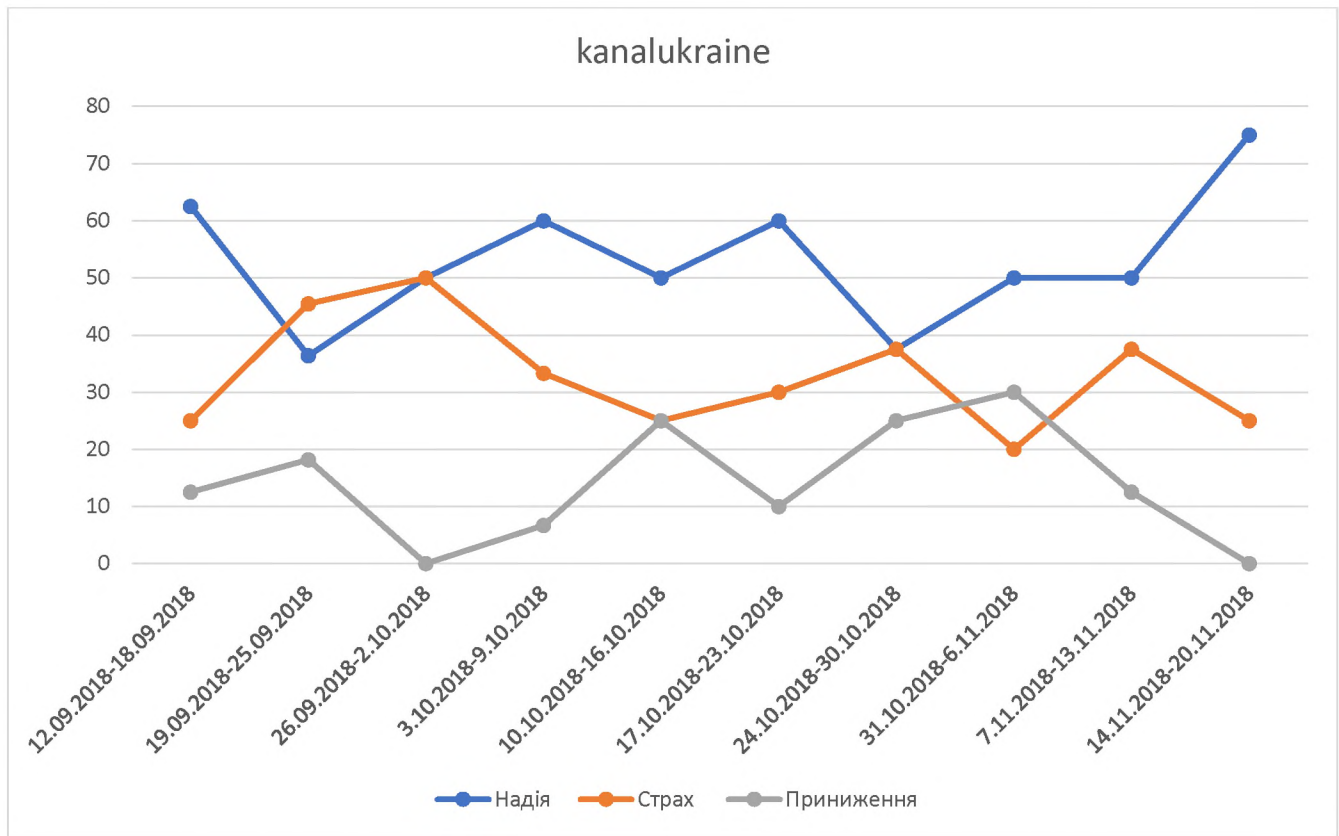


Рисунок 3.4 – Часовий ряд за даними групи kanalukraine

### 3.3.5 Результат групи News On Ukraine

Результати проведення класифікації твітів групи News On Ukraine представлені у Таблиці 3.5. А саме кількість твітів з емоцією надії, страху та приниження та загальною кількістю твітів. Дані розбиті на 10 тижнів починаючи з 12.09.2018 по 20.11.2018.

Таблиця 3.5 – Класифіковані дані групи News On Ukraine

News On Ukraine				
	Надія	Страх	Приниження	Всього
12.09.2018-18.09.2018	7	5	4	16
19.09.2018-25.09.2018	8	4	1	13
26.09.2018-2.10.2018	3	4	3	10
3.10.2018-9.10.2018	5	7	3	15
10.10.2018-16.10.2018	15	5	1	21
17.10.2018-23.10.2018	3	1	1	5
24.10.2018-30.10.2018	2	0	0	2
31.10.2018-6.11.2018	2	1	0	3
7.11.2018-13.11.2018	2	2	1	5
14.11.2018-20.11.2018	1	1	0	2

За отриманими результатами було побудовано часовий ряд який представлений на Рисунку 3.5.

Були визначенні функції ліній тренду у поліноміальній формі 2-го ступеню.

$$y = -0,9727x^2 + 12,381x + 25,025 \text{ – для емоції надії}$$

$$y = 0,9109x^2 - 9,6787x + 49,748 \text{ – для емоції страху}$$

$$y = 0,0619x^2 - 2,7026x + 25,227 \text{ – для емоції приниження}$$

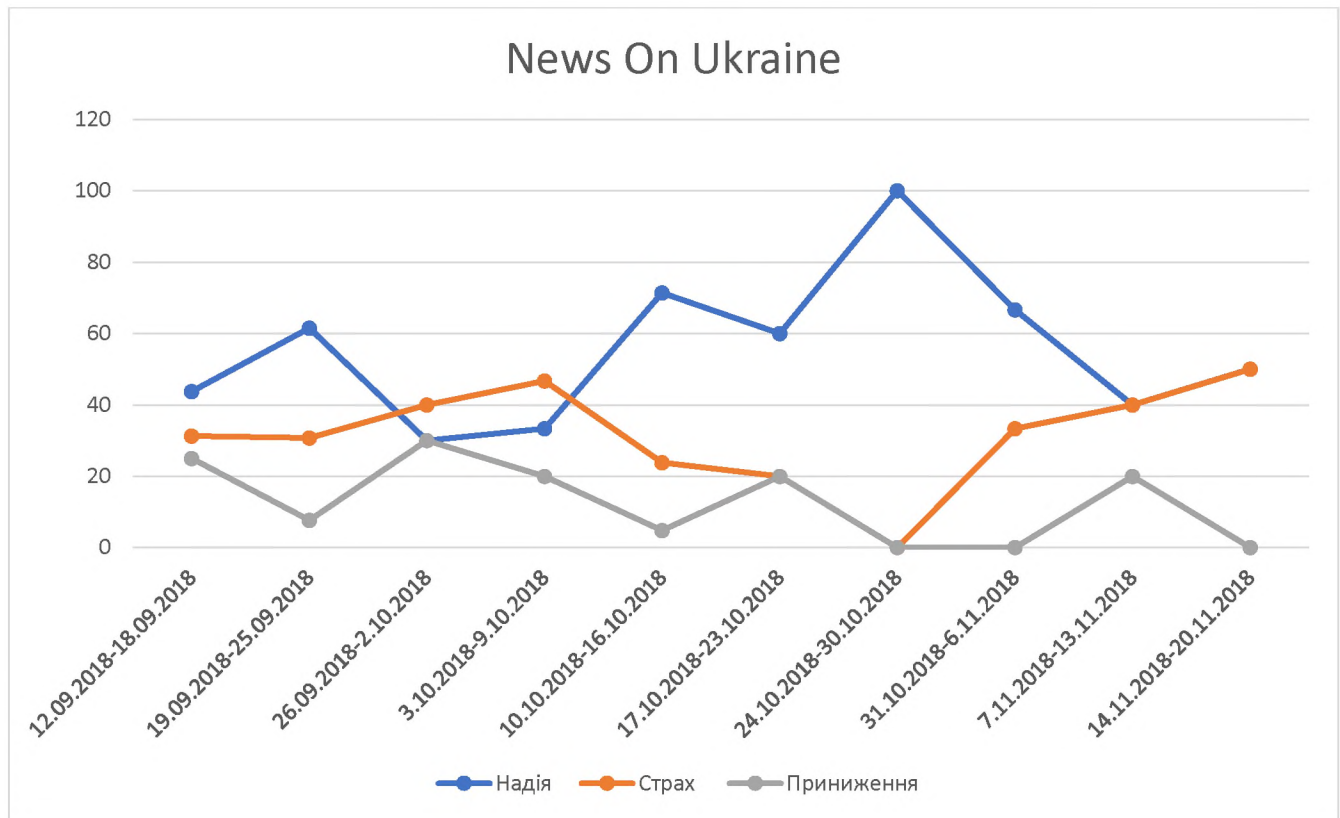


Рисунок 3.5 – Часовий ряд за даними групи News On Ukraine

### 3.3.6 Результат за 5 групами

Результати проведення класифікації твітів за 5 групами представленні у Таблиці 3.6. А саме кількість твітів з емоцією надії, страху та приниження та загальною кількістю твітів. Дані розбиті на 10 тижнів починаючи з 12.09.2018 по 20.11.2018.

Таблиця 3.6 – Класифіковані дані за 5 групами

Загальна за 5 груп				
	Надія	Страх	Приниження	Всього
12.09.2018-18.09.2018	160	135	23	318
19.09.2018-25.09.2018	138	137	26	301
26.09.2018-2.10.2018	122	139	23	284
3.10.2018-9.10.2018	160	123	35	318
10.10.2018-16.10.2018	141	82	33	256
17.10.2018-23.10.2018	142	128	42	312
24.10.2018-30.10.2018	152	118	36	306
31.10.2018-6.11.2018	166	163	33	362
7.11.2018-13.11.2018	150	162	38	350
14.11.2018-20.11.2018	105	88	25	218

За отриманими результатами було побудовано часовий ряд який представлений на Рисунку 3.6.

Були визначенні функції ліній тренду у поліноміальній формі 2-го ступеню.

$$y = -0,0866x^2 + 0,7264x + 46,995 \text{ – для емоції надії}$$

$$y = 0,2371x^2 - 2,7556x + 47,916 \text{ – для емоції страху}$$

$$y = -0,1505x^2 + 2,0292x + 5,0886 \text{ – для емоції приниження}$$

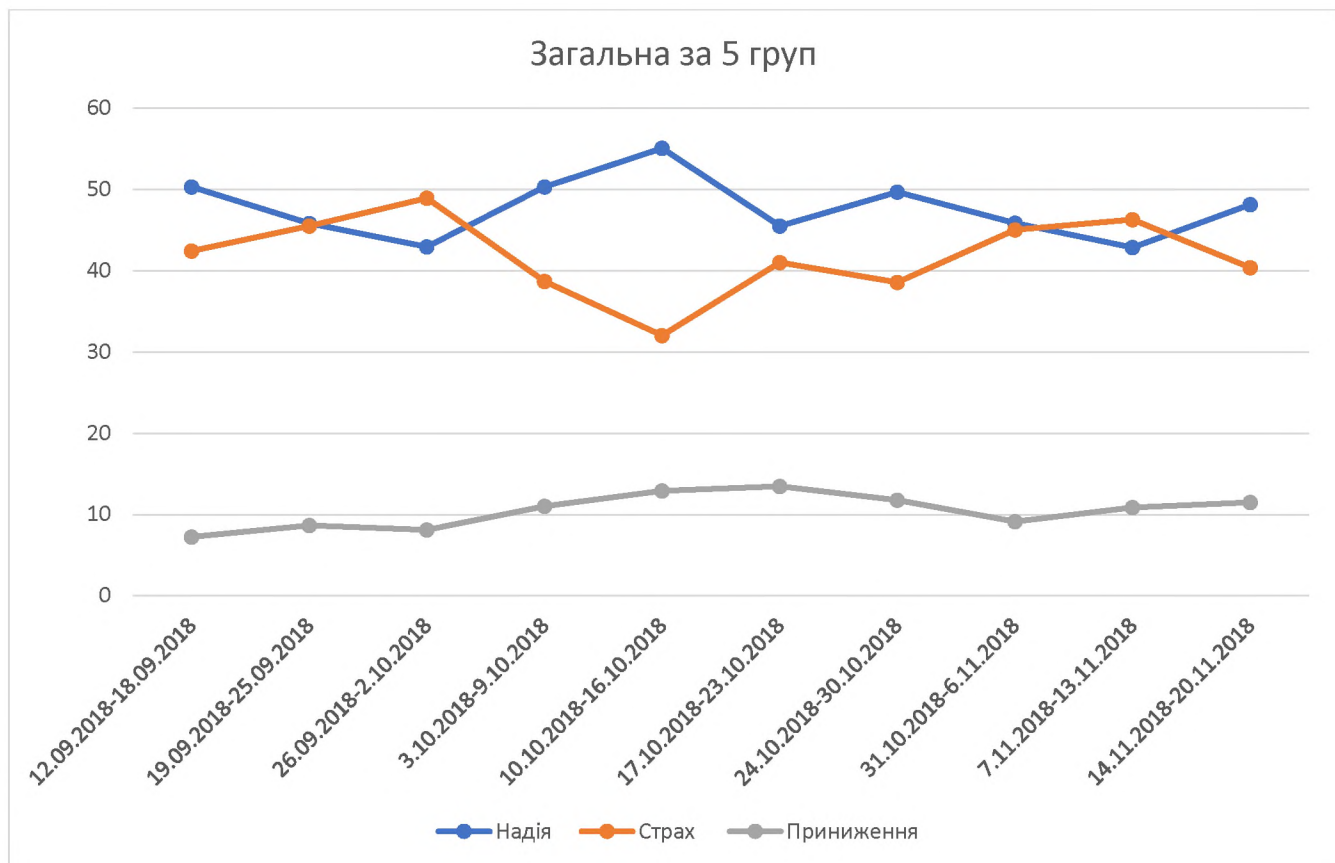


Рисунок 3.6 – Часовий ряд за даними 5 груп

### Висновки до розділу 3

Були зібранні данні з соціальної мережі Twitter. Для збору даних було вибрано 5 сторінок в Twitter.

Для класифікації твітів була створена нейрона мережа.

Отримані класифіковані дані представлені у таблицях. Та за отримані результати були представленні у вигляді часових рядів, які складаються з 10 тижнів.

## ВИСНОВКИ

За результатами виконання дипломної роботи була досягнута поставлена мета. А саме була отримана оцінка емоційного стану суспільства використовуючи соціальні медіа.

Була написана програма, яка взаємодіє з Twitter API для завантаження даних з соціальної мережі.

Для класифікації отриманих твітів була створена нейронна мережа на базі відкритої бібліотеки TensorFlow. Точність класифікації нейронної мережі дорівнює 65%

В результаті були отримані дані кількості твітів для 5 груп з новинами з емоціями:

1. Надії
2. Страху
3. Припинення

За даними результатами було побудовано 5 часових рядів для кожної з груп, які оцінювались. Дані розбиті на 10 тижнів:

1. 12.09.2018- 18.09.2018
2. 19.09.2018-25.09.2018
3. 26.09.2018-2.10.2018
4. 3.10.2018-9.10.2018
5. 10.10.2018-16.10.2018
6. 17.10.2018-23.10.2018
7. 24.10.2018-30.10.2018
8. 31.10.2018-6.11.2018
9. 7.11.2018-13.11.2018
10. 14.11.2018-20.11.2018

Для кожного часового ряду були визначенні функції тренду у вигляді поліному 2-го ступеня.

Та був порахований загальний результат за 5 групами.

Для кожної з груп твітів з емоцією надії та з емоцією страху публікується приблизно однакова кількість з деякими перепадами. В той час коли твітів з емоціями приниження публікується набагато менше.

В результаті для 5 груп кількість твітів з емоцією надії коливається в межах 43%-55%, для емоції страху коливання відбувається в межах 32%-48%, та для емоції приниження коливання кількості твітів відбувається в межах 7%-13%.

Отже можна зробити висновок, що в групах в яких відбувалося спостереження домінують твіти з емоцією надії за ними йде кількість твітів з емоцією страху та малу частину становлять кількість твітів з емоцією приниження.

## ПЕРЕЛІК ДЖЕРЕЛ ПОСИЛАННЯ

1. Сущность, функции и виды чувств и эмоций [Электронный ресурс] – Режим доступа до ресурсу: <http://www.grandars.ru/college/psihologiya/emocii-i-chuvstva.html>.
2. Моизи Д. Геополитика эмоций / Доминик Моизи. – Москва: Московская школа политических исследований, 2010. – 216 с.
3. Эмоция [Электронный ресурс] – Режим доступа до ресурсу: <https://ru.wikipedia.org/wiki/Эмоция>.
4. Емоція [Электронный ресурс] – Режим доступа до ресурсу: <https://uk.wikipedia.org/wiki/Емоція>.
5. Надія [Электронный ресурс] – Режим доступа до ресурсу: <https://uk.wikipedia.org/wiki/Надія>.
6. Страх [Электронный ресурс] – Режим доступа до ресурсу: <https://uk.wikipedia.org/wiki/Страх>.
7. Бонцанини М. Анализ социальных медиа на Python / Марко Бонцанини., 2018. – 288 с.
8. Анализ данных [Электронный ресурс] – Режим доступа до ресурсу: [https://ru.wikipedia.org/wiki/Анализ\\_данных](https://ru.wikipedia.org/wiki/Анализ_данных).
9. Грас Д. Data Sciencia Наука о данных с нуля / Джоэл Грас. – Санкт-Петербург: БВХ-Петербург, 2017. – 336 с.
10. Нейронные сети для начинающих. [Электронный ресурс] – Режим доступа до ресурсу: <https://habr.com/post/312450/>.
11. Что такое свёрточная нейронная сеть [Электронный ресурс] – Режим доступа до ресурсу: <https://habr.com/post/309508/>.
12. Свёрточная нейронная сеть [Электронный ресурс] – Режим доступа до ресурсу: <https://habr.com/post/348000/>.
13. Аналіз часових рядів [Электронный ресурс] – Режим доступа до ресурсу: [https://uk.wikipedia.org/wiki/Аналіз\\_часових\\_рядів](https://uk.wikipedia.org/wiki/Аналіз_часових_рядів).



- 14.Макошев. Политическая география и геополитика [Электронный ресурс] / Макошев – Режим доступа до ресурсу: <http://txtb.ru/96/index.html>.
- 15.ШКОЛИ І ТЕЧІЇ ГЕОПОЛІТИКИ [Электронный ресурс] – Режим доступа до ресурсу: <http://studies.in.ua/ru/geopolitika-seminary/1293-shkoli-techyi-geopoltiki.html>.
- 16.Геополітика [Электронный ресурс] – Режим доступа до ресурсу: <https://uk.wikipedia.org/wiki/Геополітика>.
- 17.Глобализация [Электронный ресурс] – Режим доступа до ресурсу: <http://www.grandars.ru/college/sociologiya/posledstviya-globalizacii.html>.

**ДОДАТКИ**

## ДОДАТОК А

```

import os
import sys
from tweepy import API
from tweepy import OAuthHandler
def get_twitter_auth():
    try:
        consumer_key = os.environ['TWITTER_CONSUMER_KEY']
        consumer_secret = os.environ['TWITTER_CONSUMER_SECRET']
        access_token = os.environ['TWITTER_ACCESS_TOKEN']
        access_secret = os.environ['TWITTER_ACCESS_SECRET']
    except KeyError:
        sys.stderr.write("Error key")
        sys.exit(1)
    auth = OAuthHandler(consumer_key, consumer_secret)
    auth.set_access_token(access_token, access_secret)
    return auth
def get_twitter_client():
    auth=get_twitter_auth()
    client=API(auth)
    return client
import sys
import json
from tweepy import Cursor
from twitter_client import get_twitter_client

if __name__ == '__main__':
    user = sys.argv[1]
    client = get_twitter_client()

    fname = "user_timeline_().jsonl".format(user)

    with open(fname, 'w') as f:
        for page in Cursor(client.user_timeline,
                           screen_name=user,
                           count=200).pages(16):
            for status in page:
                f.write(json.dumps(status._json)+"\n")
import pandas as pd
import numpy as np
import tensorflow as tf
from collections import Counter

# load a csv
CSV_PATH = "1.1.csv"

```

```

dataset = tf.contrib.data.make_csv_dataset(CSV_PATH, batch_size=32)
iter = dataset.make_one_shot_iterator()
next = iter.get_next()
print(next) # next is a dict with key=columns names and value=column data
inputs, labels = next['text'], next['Labels']
train_dataset = tf.data.Dataset.from_tensor_slices(inputs)
test_dataset = tf.data.Dataset.from_tensor_slices(inputs)
vocab = Counter()

for text in train_dataset .data:
    for word in text.split(' '):
        vocab[word.lower()] += 1
for text in test_dataset .data:
    for word in text.split(' '):
        vocab[word.lower()] += 1
total_words = len(vocab)
def get_word_2_index(vocab):
    word2index = {}
    for i,word in enumerate(vocab):
        word2index[word.lower()] = i
    return word2index
word2index = get_word_2_index(vocab)
def text_to_vector(text):
    layer = np.zeros(total_words,dtype=float)
    for word in text.split(' '):
        layer[word2index[word.lower()]] += 1

    return layer

def category_to_vector(category):
    y = np.zeros((3),dtype=float)
    if category == 0:
        y[0] = 1.

```

```

elif category == 1:
    y[1] = 1.
else:
    y[2] = 1.
return y
def get_batch(df,i,batch_size):
    batches = []
    results = []
    texts = df.data[i*batch_size:i*batch_size+batch_size]
    categories = df.target[i*batch_size:i*batch_size+batch_size]
    for text in texts:
        layer = text_to_vector(text)
        batches.append(layer)
    for category in categories:
        y = category_to_vector(category)
        results.append(y)
    return np.array(batches),np.array(results)
learning_rate = 0.01
training_epochs = 10
batch_size = 150
display_step = 1
# Network Parameters
n_hidden_1 = 10    # 1st layer number of features
n_hidden_2 = 10    # 2nd layer number of features
n_input = total_words # Words in vocab
n_classes = 3      # Categories: graphics, sci.space and baseball
input_tensor = tf.placeholder(tf.float32,[None, n_input],name="input")
output_tensor = tf.placeholder(tf.float32,[None, n_classes],name="output")
def multilayer_perceptron(input_tensor, weights, biases):
    layer_1_multiplication = tf.matmul(input_tensor, weights['h1'])
    layer_1_addition = tf.add(layer_1_multiplication, biases['b1'])
    layer_1 = tf.nn.relu(layer_1_addition)

```

```

# Hidden layer with RELU activation
layer_2_multiplication = tf.matmul(layer_1, weights['h2'])
layer_2_addition = tf.add(layer_2_multiplication, biases['b2'])
layer_2 = tf.nn.relu(layer_2_addition)

# Output layer
out_layer_multiplication = tf.matmul(layer_2, weights['out'])
out_layer_addition = out_layer_multiplication + biases['out']

return out_layer_addition

# Store layers weight & bias
weights = {
    'h1': tf.Variable(tf.random_normal([n_input, n_hidden_1])),
    'h2': tf.Variable(tf.random_normal([n_hidden_1, n_hidden_2])),
    'out': tf.Variable(tf.random_normal([n_hidden_2, n_classes]))
}
biases = {
    'b1': tf.Variable(tf.random_normal([n_hidden_1])),
    'b2': tf.Variable(tf.random_normal([n_hidden_2])),
    'out': tf.Variable(tf.random_normal([n_classes]))
}

# Construct model
prediction = multilayer_perceptron(input_tensor, weights, biases)

# Define loss and optimizer
loss = tf.reduce_mean(tf.nn.softmax_cross_entropy_with_logits(logits=prediction,
labels=output_tensor))

optimizer = tf.train.AdamOptimizer(learning_rate=learning_rate).minimize(loss)

# Initializing the variables
init = tf.global_variables_initializer()

saver = tf.train.Saver()

```

with tf.Session() as sess:

```

    sess.run(init)
    for epoch in range(training_epochs):
        avg_cost = 0.
        total_batch = int(len(train_dataset.data)/batch_size)
        # Loop over all batches
        for i in range(total_batch):
            batch_x, batch_y = get_batch(train_dataset, i, batch_size)
            # Run optimization op (backprop) and cost op (to get loss value)
            c, _ = sess.run([loss, optimizer], feed_dict={input_tensor:
batch_x, output_tensor: batch_y})
            # Compute average loss
            avg_cost += c / total_batch
        # Display logs per epoch step
        if epoch % display_step == 0:
            print("Epoch:", '%04d' % (epoch+1), "loss=", \
                  "{:.9f}".format(avg_cost))
        correct_prediction = tf.equal(tf.argmax(prediction, 1), tf.argmax(output_tensor, 1))
        # Calculate accuracy
        accuracy = tf.reduce_mean(tf.cast(correct_prediction, "float"))
        total_test_data = len(test_dataset.target)
        batch_x_test, batch_y_test = get_batch(test_dataset, 0, total_test_data)
        print("Accuracy:", accuracy.eval({input_tensor: batch_x_test, output_tensor: batch_y_test}))

```

## ДОДАТОК В

{"text": "UK intel ex-operative reveals Russia's cunning tactics in hybrid war against Ukraine\nhttps://t.co/R0ehGmVOtl https://t.co/cZnjVnzbR6"}	0
{"text": "Ryanair to launch five new flights from Ukraine in summer\nhttps://t.co/tj5M8a5oeS https://t.co/oG3voRIv1y"}	0
{"text": "Poroshenko appoints new governor in Cherkasy region\nhttps://t.co/EorR2HUTl1 https://t.co/f4HP7XePHG"}	0
{"text": "TYME international payment system files lawsuit against NBU\nhttps://t.co/nKZLwYQwx4 https://t.co/ufOOAtzabz"}	1
{"text": "Klitschko vows to buy cable cars with air conditioning for Kyiv funicular railway in 2019\nhttps://t.co/9O19mNoQzO https://t.co/yMgAiDBcmm"}	0
{"text": "Poroshenko: NATO's, EU's doors open for Ukraine\nhttps://t.co/YioweBnxKL https://t.co/MVxMYqSO10"}	0
{"text": "Ukraine may start economic blockade of Crimea in response to Russia's threats in Sea of Azov \u2013 https://t.co/o3jvi7byuB"}	1
{"text": "EU cheerleaders go to Russia-occupied Ukraine \u2013 media\nhttps://t.co/kdUSE0mjJI https://t.co/ut5WSF3lag"}	2
{"text": "Ukrtsukor revises downward sugar production forecast for 2018-2019 MY\nhttps://t.co/xhZLgSkIxT https://t.co/OyDXO4LPwX"}	1
{"text": "Owners of imported cars with foreign registration block Kyiv-Chop highway in Lviv region\nhttps://t.co/cdW1HiwTk8 https://t.co/ieLRsbw4km"}	1
{"text": "Brother of Russia's candidate for Interpol chief represents Ukraine at OSCE \u2013 media\nhttps://t.co/ooGSS14INx https://t.co/1ylJxgcpEm"}	1
{"text": "Another five U.S.-built locomotives arrive in Ukraine\nhttps://t.co/HpfppEwxOe https://t.co/6ZvLnRfenR"}	0
{"text": "EU mulls new Russia sanctions over Ukraine 'elections' \u2013 media\nhttps://t.co/UEeogNXtlW https://t.co/8TACrfgZDJ"}	0
{"text": "Deaths in Ukraine almost twice birth rate in nine months (Document)\nhttps://t.co/odSEVFKin5 https://t.co/NRAc6giNWZ"}	1
{"text": "Some 60 towns and villages in Ukraine see power outages due to bad weather\nhttps://t.co/FpdUQSqzpj https://t.co/U1Tn6fsN1V"}	1
{"text": "Ukraine's air defense unit shoots down Russian Orlan-10 drone in Donbas (Photo)\nhttps://t.co/u8XE8t64Zt https://t.co/sWPAPVXRkc"}	1
{"text": "Owners of imported cars with foreign registration blocking traffic in Kyiv (Video)\nhttps://t.co/chab81sXrN https://t.co/2SrksBIMXh"}	2
{"text": "Putin mulls Russia retaliation if U.S. quits INF Treaty \u2013 media\nhttps://t.co/lmkJ0mMpzm https://t.co/aLQ7iu0R70"}	1
{"text": "Three Ukrainian soldiers wounded in Donbas in past 24 hours\nhttps://t.co/A8UDvIXSJx https://t.co/Z9ue3sJJmF"}	1
{"text": "NBU cuts official forex rate to UAH 27.76 to dollar\nhttps://t.co/oBNzF8Ys61 https://t.co/RVyLAhTUd3"}	0



{"text": "Reuters: Oil prices drop as broad market sell-off outweighs supply concerns\nhttps://t.co/cKVQARFgwA https://t.co/7HHI01YOPa"} 0  
 {"text": "EU Council ready to introduce additional sanctions over pseudo-elections in Donbas \u2013\nhttps://t.co/iERfmWXgxi https://t.co/zt3YkKJzul"} 0  
 {"text": "Hungary to continue blocking Ukraine-NATO Commission\nhttps://t.co/M6H4Sp9Tjp https://t.co/e12voWuCVp"} 2  
 {"text": "Shale gas production could increase Ukraine's resource base fivefold \u2013\nUkr gazvydobuvannya\nhttps://t.co/yakwSwmy2D https://t.co/83XflGewR4"} 0  
 {"text": "Putin's people to face failure at Ukraine elections \u2013\nexpert\nhttps://t.co/GoWY5VR2ok https://t.co/WpL98AbAs5"} 0  
 {"text": "Miners in \"LPR\" protesting against wage arrears \"stripped, forced to dig holes\"\nhttps://t.co/UwIb5cxYPh https://t.co/bocjZw7l26"} 2  
 {"text": "EU reaffirms non-recognition of sham \"elections\" in Donbas\nhttps://t.co/sLaeSnpC2M https://t.co/soshtlkbOY"} 0  
 {"text": "Russia detains Ukrainian fishing vessel with crew in Sea of Azov - media\nhttps://t.co/KHAGiL1EFO https://t.co/P2dy0th0Fc"} 1  
 {"text": "Ukraine official condemns sale of products with \"DPR\" symbols on Amazon\nhttps://t.co/YH6MyZ2xW0 https://t.co/iGeCyAfn4I"} 0  
 {"text": "Erdogan, Putin mark completion of offshore section of Turkish Stream\nhttps://t.co/IWVGOUVObk https://t.co/Ofs8V5xywc"} 1  
 {"text": "How is the new UN General Assembly resolution on Crimea important for Ukraine\nhttps://t.co/02t0jln9G https://t.co/YIT96lKVC"} 0  
 {"text": "Ukraine may suspend membership in Interpol - Interior Minister Avakov\nhttps://t.co/Pb6iqdKsVY https://t.co/Lccr81tyRq"} 2  
 {"text": "Kobolyev considers Naftogaz's unbundling as condition for further gas transit via Ukraine\nhttps://t.co/QMBFVMhME0 https://t.co/LdjBRdBkW3"} 2  
 {"text": "Many Russian servicemen reportedly refuse to serve in Donbas \u2013\nUkraine's military intel\nhttps://t.co/0c9XPtFWDt https://t.co/GdJiAzzkkx"} 0  
 {"text": "National Bank forecasts levels of Ukrainians' salaries in coming two years\nhttps://t.co/a2aP4jiJRT https://t.co/7uAXb8jFgU"} 0  
 {"text": "Trial of Ukraine's ex-president Yanukovich postponed due to his hospitalization \u2013\nmedia\nhttps://t.co/y39r7fbzG8 https://t.co/As7qa3QoS"} 2  
 {"text": "Ukraine manufactures prototype of APC hull from steel supplied by NATO ally\nhttps://t.co/fNa7ZUR6XT https://t.co/itC2Xmn7K"} 0  
 {"text": "\"DPR parliament\" headed by former Communist MP (Photo)\nhttps://t.co/p2wQGGbRV5 https://t.co/P5I4TiVsw"} 2  
 {"text": "Ukrainian gas stocks shrink by 1.7% to 16.9 bcm since Nov 7\nhttps://t.co/wZ30bKsWBb https://t.co/mdbJfaGbNH"} 2  
 {"text": "Patriarch Filaret comments on his possible election as head of Ukraine's local Orthodox church\nhttps://t.co/DXW4uloeUP"} 0

{"text": "Ukraine cuts grain exports by 12% in Jan-Sept \u2013 expert\nhttps://t.co/0xoSDOgO4Q https://t.co/Tb0joJ7mo8"}	1
{"text": "PM Groyzman: Ukraine needs external support to pay off previous debts\nhttps://t.co/8WuhaWTShX https://t.co/EhguuRoOYU"}	2
{"text": "Russia spends about US\$2 bln per year on \"LPR/DPR\" \u2013 Putin's ex-advisor\nhttps://t.co/7TJ7O5hUIQ https://t.co/Ja6jynbIIC"}	2
{"text": "U.S., Ukraine in 'close discussion' for new lethal arms \u2013 media\nhttps://t.co/6vv3DwQr5b https://t.co/x6vF12SQW1"}	0
{"text": "Economist forecasts UAH 30 per U.S. dollar in Ukraine by end of 2018\nhttps://t.co/VIv12y506V https://t.co/BfK4dp2hZ5"}	1
{"text": "Holy Council of Ukraine's Orthodox Church scheduled for December \u2013 Ecumenical Patriarchate\nhttps://t.co/SmSIADCXUE https://t.co/dpBgy0CG6F"}	0
{"text": "Political analyst explains why Armenian church opposes Ukrainian autocephaly\nhttps://t.co/mu8eBcq5lO https://t.co/PEoUPllBDd"}	1
{"text": "Over 44,000 people cross border at checkpoints with occupied Donbas within 24 hours\nhttps://t.co/I0dPNOzThM https://t.co/OJt9NNFtcx"}	2
{"text": "Expert warns about Russia's provocations before church unification council\nhttps://t.co/yJEh4MVP9D https://t.co/KkaJlW5var"}	1
{"text": "Damp day of protests in Sofia, other Bulgarian cities \u2013 media\nhttps://t.co/PdXKAw290a https://t.co/PbpVI4Up10"}	1
{"text": "Ukraine reports 16 enemy attacks in Donbas over past 24 hours\nhttps://t.co/pk4YZ2aJ9N https://t.co/ihrOCaA2jR"}	1
{"text": "Reuters: Oil rises on expected OPEC cuts, but markets wary on trade woes\nhttps://t.co/LrvRr5tpUW https://t.co/8vKvj1SDbC"}	1
{"text": "JFO: Ukraine reports eight enemy attacks in Donbas on Nov 18\nhttps://t.co/kDPYXJHwIA https://t.co/muLfh1X2uK"}	1
{"text": "Unification council of Ukrainian Orthodox churches postponed for December - media\nhttps://t.co/v5zUGmVc6k https://t.co/2T11EDXJ9h"}	0
{"text": "OSCE spots amassing of heavy weapons in occupied Donbas\nhttps://t.co/kxNuRniyDd https://t.co/t0bAjPKJjn"}	0
{"text": "About 12,000 Ukrainians recruited into Russian army since Crimea occupation \u2013 human rights...\u2013 https://t.co/qpbWM5jlHa"}	2
{"text": "RFE/RL: Kyiv march for transgender rights cancelled after violence by far-right radicals\nhttps://t.co/Ek0DI6CjZx https://t.co/sDB04gp2FA"}	0
{"text": "Ukrainian FM Klimkin: Russia loses war in Ukraine\nhttps://t.co/trMK2RpaZr https://t.co/U96tXHPp42"}	1
{"text": "Ukraine-China trade sees huge rise \u2013 media\nhttps://t.co/wyWCzH0UcC https://t.co/4LXfA7jFpi"}	0
{"text": "Kyiv Patriarchate refutes media reports on holding unification council on Nov 22\nhttps://t.co/BSXQwJiP3m https://t.co/C4j0Nid0HH"}	0
{"text": "Two killed after vintage plane crashes in U.S. \u2013 media\nhttps://t.co/l7PwMxEbFa https://t.co/WE7oAiuYru"}	1
{"text": "Yanukovich reportedly hospitalized in Moscow \u2013 Russian media\nhttps://t.co/gIqFlNa5Wq https://t.co/HLVJzeIhdv"}	0

{"text": "NBU strengthens official forex rate to UAH 27.75 to dollar for Nov 19\nhttps://t.co/QMJEN66nAW https://t.co/vt7p82qAQ3"} 0  
 {"text": "Donbas update: One Ukrainian soldier killed in past 24 hours\nhttps://t.co/RcrHzKExpV https://t.co/tlQ4ljbEYU"} 1  
 {"text": "MEPs, experts urge Facebook to open office in Kyiv to counter disinformation during Ukraine...\u2026 https://t.co/v1VWs5sFGj"} 0  
 {"text": "Ukraine's defense minister responds to Russian threats to \"shut down\" Sea of Azov\nhttps://t.co/80T9pWCIsW https://t.co/pWM56jrDhK"} 1  
 {"text": "Belarus Putin's No. 1 target \u2013 ex-advisor\nhttps://t.co/SVjMw2Upbu https://t.co/qIeLQ6TK1U"} 1  
 {"text": "Week's balance: Poroshenko \"turns on\" heating, U.S. vows energy support, while GDP growth...\nhttps://t.co/m1QhzAIEZg https://t.co/LuHISil0TE"} 0  
 {"text": "Russian official may become Interpol's next president\nhttps://t.co/Y1f0rXT8O7 https://t.co/PCeXX7U0i7"} 1  
 {"text": "U.S. to continue to impose Russia sanctions until Crimea, Donbas returned to Ukraine - Pompeo\u2026 https://t.co/bRw1Bfn6vB"} 0  
 {"text": "Putin's ex-adviser comments on Russia's possible reaction to Ukraine church independence\nhttps://t.co/gD3rKcS2vw https://t.co/f4XTpfD0yY"} 2  
 {"text": "OSINT group elaborates on Russia's \"active measures\" in Ukraine\nhttps://t.co/rH0MLplBeb https://t.co/wBTkIHMDdN"} 0  
 {"text": "EU may impose sanctions against Hungary \u2013 MP Arie\u00f1ev\nhttps://t.co/NZA7lWsbRS https://t.co/nDWv4Wicil"} 0  
 {"text": "U.S. to help Ukraine fight Russian meddling in 2019 elections\nhttps://t.co/z9yWqx2jQP https://t.co/HzuID7guMw"} 0  
 {"text": "Activists rally in Kyiv demanding interior minister Avakov's resignation (Photo, video)\nhttps://t.co/SslSdsvJ2N https://t.co/4psttwvkBg"} 1  
 {"text": "U.S. condemns Russia's aggressive actions in Azov-Black Sea region.\nhttps://t.co/2IQo3VDLkE https://t.co/bmK19LzPaV"} 0  
 {"text": "Medvedchuk not Putin's agent but \"direct representative\" in Ukraine \u2013 Ex PM\nhttps://t.co/bxlGdedTyE https://t.co/KOzk8d69Hb"} 2  
 {"text": "Bank collectors robbed outside Kyiv\nhttps://t.co/HFEgXOiAFW https://t.co/SdwPt3CKHv"} 1  
 {"text": "Ukraine's Foreign Intelligence Service helps thwart another massive cyber attack on gov't...\nhttps://t.co/6syaH62YHK https://t.co/1LePfYhXQ4"} 0  
 {"text": "Another Ukrainian soldier killed in Donbas as fighting continues\nhttps://t.co/Ng03To4PyJ https://t.co/OJjBUNC2BF"} 1  
 {"text": "CIA has concluded Saudi crown prince ordered journalist's killing: Washington Post\nhttps://t.co/4xuq3zNT0V https://t.co/bY2U6qI499"} 1  
 {"text": "Ukraine, U.S. create three new bilateral working groups: Security, rule of law, economy\nhttps://t.co/5yLQQVVXT2 https://t.co/oaME1j3oJd"} 0  
 {"text": "Pompeo to Klimkin: We'll keep working together to stop Nord Stream 2 (Video)\nhttps://t.co/EaX6E5Tzem https://t.co/5CcOy7OdWN"} 0  
 {"text": "#US #Ukraine #partnership https://t.co/j0rt3PN1OH"} 0

{"text": "New Hungarian ambassador summoned to Ukraine's Foreign Ministry over Orban's statement\nhttps://t.co/dXb7T6kaJC https://t.co/SDYzKrhDXd"} 0

{"text": "Klimkin: Ukraine, U.S. discuss cooperation in arms supplies (Video)\nhttps://t.co/RhaoVhf1VO https://t.co/HtiHlweyBJ"} 0

{"text": "BBC: Russia 'sought access to UK visa issuing system'\nhttps://t.co/sxITnT3IaX https://t.co/1f9UH4RPR3"} 1

{"text": "Part of sidewalk of bridge near Kyiv's Darnytsia subway station reportedly collapses (Photo,...\u2026 https://t.co/MbHpBJYA12"} 1

{"text": "Russian-led forces mount five attacks on Ukraine's army in Donbas on Friday, no casualties...\u2026 https://t.co/HPcAQJ52dd"} 1

{"text": "Poroshenko enacts agreement on EUR 1 bln macro-financial assistance from EU\nhttps://t.co/0m9LDEpKcd https://t.co/Gtj4Y8WEDV"} 0

{"text": "Ex-adviser to Putin explains why Kremlin needs \"LPR/DPR\"'\nhttps://t.co/XXNwtVgHYG https://t.co/23LPdPzHVF"} 2

{"text": "Turchynov rules out any talks with \"DPR/LPR\"'\nhttps://t.co/Q7f9BtLlfU https://t.co/DRuNUHglPE"} 0

{"text": "Ukraine may boost agricultural output by 5.1% in 2018 \u2013 experts\nhttps://t.co/u5EallnTcA https://t.co/HZJqxMs8ED"} 0

{"text": "Cost of Ukraine's e-visas goes up\nhttps://t.co/bwWwf7y4RN https://t.co/lkchNoyIrb"} 2

{"text": "Ukraine's Minister Zhdanov refutes foreign media allegations about \"nationalist camp training...\u2026 https://t.co/glza4n8enW"} 2

{"text": "Russia redirects part of naval forces near Crimea to new bridgehead \u2013 2013 border guards\nhttps://t.co/DkKfzzZq3F https://t.co/kuTRfRZdMT"} 1

{"text": "Ukraine to buy liquefied gas, coal until it covers shortage with own resources \u2013 PM Groysman\nhttps://t.co/rfhd4zBQri https://t.co/EjFqiHugo2"} 1

{"text": "Theresa May fights to retain post as no-confidence bid grows - media\nhttps://t.co/HB3NNhAqeC https://t.co/ZwGvOnO6PO"} 0

{"text": "SAPO issues suspicion notice to ex-fiscal service head Prodan on grand embezzlement charges\nhttps://t.co/Ty7z33Uvzr https://t.co/IdnSDL0xqO"} 0

{"text": "Journalistic investigation reveals how \"Ukrainian experts\" appear on Russian propaganda TV...\u2026 https://t.co/d0aAaFslK6"} 0

{"text": "Russia's FSB conducting searches at homes of Jehovah's Witnesses in Crimea \u2013 2013 journalist\nhttps://t.co/Wcp3aWpHvy https://t.co/JmK5JDBcJG"} 1

{"text": "Ukrainian border guard in coma following hit-and-run incident involving suspected smugglers\nhttps://t.co/FyivOoUI7v https://t.co/mB5M7fUx7G"} 1

{"text": "Economy ministry to oblige all businesses to accept non-cash payments for goods, services\nhttps://t.co/A3JfIRPNC0 https://t.co/QsTLPowIA8"} 0

{"text": "Poroshenko thanks UN Member States for new resolution on Crimea\nhttps://t.co/WIHwlMDCDj https://t.co/bEyJr7sTPI"}	0
{"text": "Hungary's Orban echoes Putin, claiming no solution can be reached with today's Ukrainian...\nhttps://t.co/xDUlrlx4ow https://t.co/Giw10nBZhS"}	2
{"text": "Infrastructure minister vows US\$3 bln investment in Ukrainian sea ports over 5 years\nhttps://t.co/yBlMZ3jnt0 https://t.co/l3ld5SeYrU"}	0
{"text": "Putin's statement example of meddling in Ukraine election campaign \u2013 Poroshenko's spox\nhttps://t.co/5huWcBIOtL https://t.co/QdlBaleg5X"}	1
{"text": "Russia unable to bring Ukrainians to their knees - Klimkin\nhttps://t.co/jhZnzJmXKH https://t.co/fOnlyqgwiJ"}	0
{"text": "Ukraine's Pavlo Hryb refuses to testify at Russia trial\nhttps://t.co/dcbBUxtx5q https://t.co/frUYNLWpHn"}	0
{"text": "Russian economist calculates Kremlin's losses caused by sanctions\nhttps://t.co/D545yxSkMk https://t.co/5WWYgjK2wD"}	0
{"text": "Ukraine issues unique hexagonal coins (Photo)\nhttps://t.co/c1uIGsUYum https://t.co/Nx0NajExaN"}	0
{"text": "Two Ukrainian troops wounded in Donbas in past 24 hours\nhttps://t.co/6r0T8fJeGT https://t.co/JFcPdf4M13"}	1
{"text": "Klimkin, Volker discuss ways to counter Russia's attempts to legitimize occupation of Donbas,...\u2013 https://t.co/KOcJjcQuJW"}	0
{"text": "Russia threatens to \"shut down\" Azov Sea for Ukrainian ships \"within minutes\"\nhttps://t.co/iLjicQQWco https://t.co/hJwd1cuLte"}	1
{"text": "UNGA's 3rd Committee endorses updated draft resolution on human rights in occupied Crimea\nhttps://t.co/3dqSKINOWb https://t.co/R6gzYc23P8"}	0
{"text": "VOA: Ukraine PM upbeat on IMF loan prospects\nhttps://t.co/LtMOGOQF37 https://t.co/6tDpSdevGr"}	2
{"text": "Czechs protest against PM Babis, coalition partner may quit government \u2013 media\nhttps://t.co/Pvn4pDhx7l https://t.co/colfczqjD7"}	1
{"text": "Naftogaz postpones eurobond issue because of market volatility\nhttps://t.co/UbfdOLmvYd https://t.co/i3QMAeSyGB"}	1
{"text": "IMF expects Ukrainian authorities will increase heating tariffs\nhttps://t.co/KC9qAqvzeP https://t.co/ssLqUigfxi"}	2
{"text": "Reuters: U.S. sanctions 17 Saudis over killing of journalist Khashoggi\nhttps://t.co/5d6wMPN5IG https://t.co/cy1FSoKiwP"}	0
{"text": "EU calls on Russia and Ukraine to pursue detainees swap by year-end\nhttps://t.co/2OOEUkAJo5 https://t.co/2wvt26OqAH"}	0
{"text": "EU to sanction five individuals over sham elections in Donbas \u2013 media\nhttps://t.co/4tQzfKGdYN https://t.co/QolhvaC2QY"}	0
{"text": "Russia uses Donbas as testing ground for new arms, disposal of old ammunition \u2013 Ukraine's...\nhttps://t.co/Eo6Lt4YLFF https://t.co/kK2Ba0zExB"}	1
{"text": "Ukraine may boost agri-exports to record US\$18 bln in 2018 \u2013 UCAB\nhttps://t.co/lk8Jz4mpOx https://t.co/PmymeBASPG"}	0

